nature

www.nature.com/nature

Vol 455 | Issue no. 7212 | 25 September 2008

America's fresh start

The next US president will lead the country back onto the world stage in many arenas, including science.

Most of America's leading scientific establishment spent this spring agitating for a 'science debate' between the presidential candidates. Not surprisingly, the debate never happened; science has rarely, if ever, been a major issue in US presidential campaigns. But its supporters shouldn't lament that fact. Many of the issues that are central to the current campaign have a strong science component (see page 442).

The most obvious example is climate change. No matter who wins, the current administration's eight-year-long pattern of denial and foot-dragging will end. Both McCain and Obama have pledged to regulate the country's greenhouse gases via mandatory emissions limits and a cap-and-trade system. Given the ongoing energy crisis, either of them may very well begin this process — Congress willing — as soon as they have taken office.

Another much-discussed topic on the campaign trail is innovation. Americans increasingly sense that their country is losing its competitive edge. Both McCain and Obama regularly talk about how re-investing in fundamental research can stimulate homegrown breakthroughs and bolster the flagging economy. And both candidates seem to be at least moderately sincere in their pledges to improve funding for the nation's basic-science agencies — although how those promises will translate to reality remains to be seen, especially as McCain has talked about a year-long freeze on all domestic discretionary spending, which would include science.

The outlook for other areas of science is even less clear. In the case of stem-cell research, Obama has vowed to lift the Bush administration's restrictions on federally funded research into human embryonic stem cells. McCain, too, voted to end such restrictions when the issue arose in the Senate, but in the heat of the campaign, he now refuses to say whether he would lift the ban as president. Instead, he emphasizes research on adult or induced pluripotent stem cells in place of human embryonic ones. This could spell dire news for the country's stem-cell biologists, many of whom have been driven overseas or into privately or state-funded parallel enterprises to continue their work.

The most worrying thing about a McCain presidency is not so much a President McCain as a Vice-President Palin. Sarah Palin, Alaska's governor and McCain's running mate, opposes all research into human embryonic stem cells. She is a creationist. And until lately, at least, she has been a sceptic of human-created climate change — a disquieting thought, as Palin recently said that energy will be "her baby" in the White House, thanks to her previous service as chair of the Alaska state oil and gas commission.

"No matter who wins, the current administration's 8-year-long pattern of denial and foot-dragging will end."

What is still unclear is how Palin's views will reconcile with McCain's. McCain has courageously bucked his party's more parochial viewpoints in the past, as when he fought for a cap-and-trade system long before it was politically popular. But his selection of Palin as a running mate suggests a new-found willingness to pander to his party's far-right wing.

Contrast that with Obama's statement on page 448, in which *Nature* asked him about the teaching of intelligent design in science classes. It is not easy to address students' questions about evolution without falling prey to the false notion of 'teaching the controversy', as the Royal Society's director of education discovered last week in a public-relations meltdown (see 'Creation and classrooms'). But Obama could not be more clear: "I do not believe it is helpful to our students to cloud discussions of science with non-scientific theories like intelligent design that are not subject to experimental scrutiny," he wrote.

The emerging economic crisis has made a trillion-dollar bailout of US banks almost certain. Thus, much as the campaign may point to improved government attitudes toward science, researchers should hold out little hope for more funds for anything but the new president's very highest priorities. ■

Creation and classrooms

Better to confront superstition with science than to disregard the superstitious.

The headlines were damning. "Leading scientist urges teaching of creationism in schools," proclaimed Britain's *The Times* newspaper on 12 September, echoing the headlines appearing that day in numerous other British media. The stories asserted that Michael Reiss, a biologist and educational researcher, an ordained Anglican minister and (at the time) the education director of the Royal Society, had explicitly advocated that state-school biology classes teach creationism.

The reports were wrong. Speaking at the British Association for the Advancement of Science's annual Festival of Science on 11 September, Reiss had articulated — as he had many times before — a view consistent with the Royal Society's official position: when students from a creationist background raise the issue in class, the teacher should explain why creationism is not science and why evolution is. Nevertheless, on 16 September the society announced Reiss's departure, arguing that the media's misinterpretation had "led to damage to the society's reputation" (see page 441).

Nature was not privy to the conversations between the reporters and editors responsible for this story, so we will leave it to them to consider how such a gross misrepresentation could have happened, and what lessons to draw from it. Nor was *Nature* privy to the Royal Society's internal deliberations about Reiss, so we will leave it to the

nature

www.nature.com/nature

Vol 455 | Issue no. 7212 | 25 September 2008

America's fresh start

The next US president will lead the country back onto the world stage in many arenas, including science.

Most of America's leading scientific establishment spent this spring agitating for a 'science debate' between the presidential candidates. Not surprisingly, the debate never happened; science has rarely, if ever, been a major issue in US presidential campaigns. But its supporters shouldn't lament that fact. Many of the issues that are central to the current campaign have a strong science component (see page 442).

The most obvious example is climate change. No matter who wins, the current administration's eight-year-long pattern of denial and foot-dragging will end. Both McCain and Obama have pledged to regulate the country's greenhouse gases via mandatory emissions limits and a cap-and-trade system. Given the ongoing energy crisis, either of them may very well begin this process — Congress willing — as soon as they have taken office.

Another much-discussed topic on the campaign trail is innovation. Americans increasingly sense that their country is losing its competitive edge. Both McCain and Obama regularly talk about how re-investing in fundamental research can stimulate homegrown breakthroughs and bolster the flagging economy. And both candidates seem to be at least moderately sincere in their pledges to improve funding for the nation's basic-science agencies — although how those promises will translate to reality remains to be seen, especially as McCain has talked about a year-long freeze on all domestic discretionary spending, which would include science.

The outlook for other areas of science is even less clear. In the case of stem-cell research, Obama has vowed to lift the Bush administration's restrictions on federally funded research into human embryonic stem cells. McCain, too, voted to end such restrictions when the issue arose in the Senate, but in the heat of the campaign, he now refuses to say whether he would lift the ban as president. Instead, he emphasizes research on adult or induced pluripotent stem cells in place of human embryonic ones. This could spell dire news for the country's stem-cell biologists, many of whom have been driven overseas or into privately or state-funded parallel enterprises to continue their work.

The most worrying thing about a McCain presidency is not so much a President McCain as a Vice-President Palin. Sarah Palin, Alaska's governor and McCain's running mate, opposes all research into human embryonic stem cells. She is a creationist. And until lately, at least, she has been a sceptic of human-created climate change — a disquieting thought, as Palin recently said that energy will be "her baby" in the White House, thanks to her previous service as chair of the Alaska state oil and gas commission.

"No matter who wins, the current administration's 8-year-long pattern of denial and foot-dragging will end."

What is still unclear is how Palin's views will reconcile with McCain's. McCain has courageously bucked his party's more parochial viewpoints in the past, as when he fought for a cap-and-trade system long before it was politically popular. But his selection of Palin as a running mate suggests a new-found willingness to pander to his party's far-right wing.

Contrast that with Obama's statement on page 448, in which *Nature* asked him about the teaching of intelligent design in science classes. It is not easy to address students' questions about evolution without falling prey to the false notion of 'teaching the controversy', as the Royal Society's director of education discovered last week in a public-relations meltdown (see 'Creation and classrooms'). But Obama could not be more clear: "I do not believe it is helpful to our students to cloud discussions of science with non-scientific theories like intelligent design that are not subject to experimental scrutiny," he wrote.

The emerging economic crisis has made a trillion-dollar bailout of US banks almost certain. Thus, much as the campaign may point to improved government attitudes toward science, researchers should hold out little hope for more funds for anything but the new president's very highest priorities. ■

Creation and classrooms

Better to confront superstition with science than to disregard the superstitious.

The headlines were damning. "Leading scientist urges teaching of creationism in schools," proclaimed Britain's *The Times* newspaper on 12 September, echoing the headlines appearing that day in numerous other British media. The stories asserted that Michael Reiss, a biologist and educational researcher, an ordained Anglican minister and (at the time) the education director of the Royal Society, had explicitly advocated that state-school biology classes teach creationism.

The reports were wrong. Speaking at the British Association for the Advancement of Science's annual Festival of Science on 11 September, Reiss had articulated — as he had many times before — a view consistent with the Royal Society's official position: when students from a creationist background raise the issue in class, the teacher should explain why creationism is not science and why evolution is. Nevertheless, on 16 September the society announced Reiss's departure, arguing that the media's misinterpretation had "led to damage to the society's reputation" (see page 441).

Nature was not privy to the conversations between the reporters and editors responsible for this story, so we will leave it to them to consider how such a gross misrepresentation could have happened, and what lessons to draw from it. Nor was *Nature* privy to the Royal Society's internal deliberations about Reiss, so we will leave it to the

officers and fellows of that body to reflect on who has done the most to damage its reputation.

The misreporting surrounding Reiss has provided a propaganda gift to creationists everywhere. So in the face of such confusion, it is encouraging to hear the unequivocal stance of one of the US presidential candidates, Barack Obama, on the issue (see page 448 and 'America's fresh start'): creationism and intelligent design should not be included in a science curriculum. But scientists and science teachers must also grapple with the central challenge that Reiss was addressing: how to respond to students who have been steeped in, or confused by, scientifically nonsensical creationist beliefs when they ask about those beliefs in science classes?

Those who argue that allowing discussion of creationism in a science class gives it legitimacy, and that students who ask about it should be firmly directed to take their questions elsewhere, are misguided.

Eugenie Scott, executive director of the National Center for Science Education in Oakland, California, and a long-time advocate for the teaching of evolution, points out that in the real world, any such shut-up-and-take-it-elsewhere response from the teacher will inevitably be perceived by the student (and his or her classmates) as a humiliating personal put-down. It will obstruct rather than encourage enquiry and understanding. It will also invite complaints from outraged parents.

What is more, it will squander what experienced educators like to call 'a teachable moment'. All too often, that moment is the one opportunity that a school has to engage resistant students and introduce them to what science has to say.

At such a moment, a much more effective approach is for the teacher to follow the route Reiss advocated: deal with the question without ridicule, but make it clear that in science, theories must be testable to be valid. 'You ask if Earth is 6,000 years old, and why the descendents of Adam and Eve have no relation to the lower animals? So how can we test those hypotheses, and what does the evidence say?'

This is a difficult and minefield-laden path for teachers to follow. For an example of just how delicate, see a 23 August report in the *New York Times* of how a teacher in Florida tackled such challenges (see http://tinyurl.com/48374f). In particular, it requires that the teachers have a confident knowledge and understanding of evolution, so that they can seize on those teachable moments competently. The sad news, according to surveys, is that too few biology teachers have such an understanding: evolution is not always taught well at the universities and colleges where teachers learn their biology. And that's in the developed world; in poor and developing countries, teachers often receive no training in evolution at all.

Biology graduates who have not encountered up-to-date evidence of evolution in action — in fossils, in microbes, in genomes — have been ill-served by their training. Higher education in general, and biology departments in particular, are at the front line of the battle between creation and evolution too. ■

Pathways to security

Self-regulation is a good first step — but synthetic-biology companies still need independent oversight.

Regulators have been slow to deal with 'dual use' biological agents such as proteins, DNA or whole organisms that are generally used for benign research, but that could also be used to inflict harm. The reasons are many — not least being the complex way in which these substances behave and interact with their environment — and the result has been a regulatory patchwork.

For example, many countries have tried to regulate the firms that produce made-to-order DNA sequences by requiring permits for export. But the paperwork required is so onerous that the companies often just discard their non-domestic orders — so information about the customers looking to acquire these sequences is lost. And oversight of domestic sales is comparatively lax.

This month, the Industry Association Synthetic Biology (IASB), a consortium of gene-synthesis companies located mainly in Europe, agreed to a series of actions that might provide a more robust solution to the bioterror problem. Several of the US companies in the market have reportedly indicated their willingness to comply. The agreement calls for member companies to develop a database of suspicious or potentially dangerous DNA sequences. The association did note the potential danger of centralizing these data, even though they are already publicly available. But the benefit, argues the IASB, is that an open-source collection will be much easier for experts to keep updated, complete and correct.

Meanwhile, the agreement calls on IASB member companies to share information about the screening processes already in use so that standard practices can be adopted. The firms have said they are willing to cooperate on this effort in a non-competitive way; the report they produced includes ideas for better policing, including a pattern-recognition approach that would be more adaptable to what most predict will be a rise in the number and variety of DNA sequences requested.

"Industry self-policing can sometimes fail dramatically."

These steps, and other proposed elements, are the beginnings of a code of conduct for the industry. The reward for this voluntary practice would be an unofficial stamp of approval from the IASB that could signal to customers (and potentially the funders of customers) that this business is worth dealing with. Although such a code of conduct is useful and welcome, compliance and enforcement will be paramount. There have been, and will probably continue to be, companies that are not interested in cooperating with any industry group, and that are happy to operate in the unregulated grey area. The ultimate hope is that customers will put economic pressure on those non-compliers to fall in line, or else lose all but the most disreputable business.

But that is just a hope. As the recent meltdowns on Wall Street have indicated, industry self-policing can sometimes fail dramatically. When bad business practices can have grave effects for the public, regulators should be firm and proactive. The IASB has taken laudable first steps in providing government regulators with guidelines they can build from. Now, the regulators need to act. ■

officers and fellows of that body to reflect on who has done the most to damage its reputation.

The misreporting surrounding Reiss has provided a propaganda gift to creationists everywhere. So in the face of such confusion, it is encouraging to hear the unequivocal stance of one of the US presidential candidates, Barack Obama, on the issue (see page 448 and 'America's fresh start'): creationism and intelligent design should not be included in a science curriculum. But scientists and science teachers must also grapple with the central challenge that Reiss was addressing: how to respond to students who have been steeped in, or confused by, scientifically nonsensical creationist beliefs when they ask about those beliefs in science classes?

Those who argue that allowing discussion of creationism in a science class gives it legitimacy, and that students who ask about it should be firmly directed to take their questions elsewhere, are misguided.

Eugenie Scott, executive director of the National Center for Science Education in Oakland, California, and a long-time advocate for the teaching of evolution, points out that in the real world, any such shut-up-and-take-it-elsewhere response from the teacher will inevitably be perceived by the student (and his or her classmates) as a humiliating personal put-down. It will obstruct rather than encourage enquiry and understanding. It will also invite complaints from outraged parents.

What is more, it will squander what experienced educators like to call 'a teachable moment'. All too often, that moment is the one opportunity that a school has to engage resistant students and introduce them to what science has to say.

At such a moment, a much more effective approach is for the teacher to follow the route Reiss advocated: deal with the question without ridicule, but make it clear that in science, theories must be testable to be valid. 'You ask if Earth is 6,000 years old, and why the descendents of Adam and Eve have no relation to the lower animals? So how can we test those hypotheses, and what does the evidence say?'

This is a difficult and minefield-laden path for teachers to follow. For an example of just how delicate, see a 23 August report in the *New York Times* of how a teacher in Florida tackled such challenges (see http://tinyurl.com/48374f). In particular, it requires that the teachers have a confident knowledge and understanding of evolution, so that they can seize on those teachable moments competently. The sad news, according to surveys, is that too few biology teachers have such an understanding: evolution is not always taught well at the universities and colleges where teachers learn their biology. And that's in the developed world; in poor and developing countries, teachers often receive no training in evolution at all.

Biology graduates who have not encountered up-to-date evidence of evolution in action — in fossils, in microbes, in genomes — have been ill-served by their training. Higher education in general, and biology departments in particular, are at the front line of the battle between creation and evolution too. ■

Pathways to security

Self-regulation is a good first step — but synthetic-biology companies still need independent oversight.

Regulators have been slow to deal with 'dual use' biological agents such as proteins, DNA or whole organisms that are generally used for benign research, but that could also be used to inflict harm. The reasons are many — not least being the complex way in which these substances behave and interact with their environment — and the result has been a regulatory patchwork.

For example, many countries have tried to regulate the firms that produce made-to-order DNA sequences by requiring permits for export. But the paperwork required is so onerous that the companies often just discard their non-domestic orders — so information about the customers looking to acquire these sequences is lost. And oversight of domestic sales is comparatively lax.

This month, the Industry Association Synthetic Biology (IASB), a consortium of gene-synthesis companies located mainly in Europe, agreed to a series of actions that might provide a more robust solution to the bioterror problem. Several of the US companies in the market have reportedly indicated their willingness to comply. The agreement calls for member companies to develop a database of suspicious or potentially dangerous DNA sequences. The association did note the potential danger of centralizing these data, even though they are already publicly available. But the benefit, argues the IASB, is that an open-source collection will be much easier for experts to keep updated, complete and correct.

Meanwhile, the agreement calls on IASB member companies to share information about the screening processes already in use so that standard practices can be adopted. The firms have said they are willing to cooperate on this effort in a non-competitive way; the report they produced includes ideas for better policing, including a pattern-recognition approach that would be more adaptable to what most predict will be a rise in the number and variety of DNA sequences requested.

"Industry self-policing can sometimes fail dramatically."

These steps, and other proposed elements, are the beginnings of a code of conduct for the industry. The reward for this voluntary practice would be an unofficial stamp of approval from the IASB that could signal to customers (and potentially the funders of customers) that this business is worth dealing with. Although such a code of conduct is useful and welcome, compliance and enforcement will be paramount. There have been, and will probably continue to be, companies that are not interested in cooperating with any industry group, and that are happy to operate in the unregulated grey area. The ultimate hope is that customers will put economic pressure on those non-compliers to fall in line, or else lose all but the most disreputable business.

But that is just a hope. As the recent meltdowns on Wall Street have indicated, industry self-policing can sometimes fail dramatically. When bad business practices can have grave effects for the public, regulators should be firm and proactive. The IASB has taken laudable first steps in providing government regulators with guidelines they can build from. Now, the regulators need to act. ■

RESEARCH HIGHLIGHTS

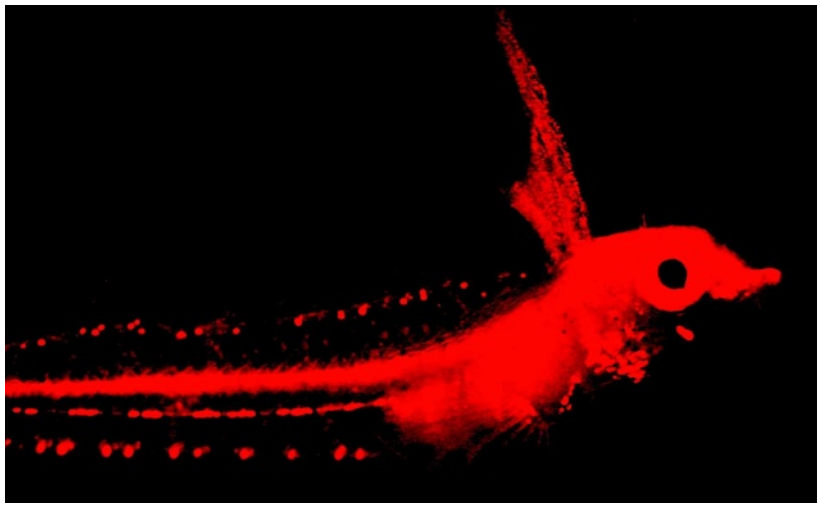
Seeing red

BMC Ecol. **8,** 16 (2008)

Water preferentially absorbs red light, so at sea it's not possible to make out red objects below a depth of around 10 metres. It was therefore thought that red was irrelevant in signalling between coral-reef fishes.

Not so. Scuba diving with filters on their masks to make red easier to pick out, Nico Michiels of the University of Tübingen in Germany and his colleagues have found at least 32 species of reef fishes that look red at depth.

These fish have crystals or pigments in their skin that fluoresce red under the incoming, mainly blue-green, light (as in the goby *Eviota destai*, pictured). The researchers say that the red markings, which are often mixed with other colours so that they appear pink, lilac or reddish brown, are most probably used for communication within each species. They add weight to this idea by showing that at least one fish, the goby *E. pellucida*, can see red.



N. K. MICHIELS *ET AL.*

CELL BIOLOGY

Home-grown fat control

Cell **134,** 933–944 (2008)

A little fat might be good for you — if it's the fatty acid palmitoleate, now discovered to be a fat-cell-derived hormone that regulates metabolism throughout the body.

Gökhan Hotamisligil at Harvard School of Public Health in Boston, Massachusetts, and his colleagues studied knockout mice lacking two proteins that normally chaperone lipids out of the bloodstream into adipose tissue. These mice are more resistant than normal mice to the bad effects of a high-fat diet. The team found increased synthesis of palmitoleate and other unsaturated fatty acids in fat cells in these mice and a consequent rise in palmitoleate levels in the blood.

They showed that palmitoleate regulates the lipid-generating program of gene expression in the liver of the mice and also stimulates muscle cells to take up glucose — and that mice infused with palmitoleate use glucose more efficiently than untreated mice.

MATERIALS SCIENCE

A rarefied insulator

Adv. Mater. **20,** 1–5 (2008)

A new material based on rare-earth elements such as cerium might overcome a barrier to making smaller silicon chips. Silicon dioxide is the traditional chip insulator, but is too bulky for smaller chips. Alternative compounds with suitably high dielectric constants are too rigid, and have proved not to insulate fully.

Now Dmitry Kukuruznyak at the Max Planck Institute for Metals Research in Stuttgart, Germany, and his colleagues have constructed another potential insulator: the rare-earth aluminium–silicon apatite $RE_6(AlO_3)_5(SiO_{3.5})$. Rather than forming rigid crystals, this compound self-organizes into flexible films on silicon at temperatures below 1,030 °C.

NONLINEAR DYNAMICS

Loading the dice

Phys. Rev. E **78,** 036207 (2008)

Gamblers — and Einstein — have assumed that throwing a die gives a random result.

But does it? Jan Nagler of the Max Planck Institute for Dynamics and Self-Organization in Göttingen and Peter Richter of the University of Bremen in Germany have simplified the throw of a die to the two-dimensional case of a dumb-bell tossed onto a surface. Will it fall with one labelled end pointing to the left or to the right?

The researchers calculated the dynamics and find that these are only truly chaotic, leading to complete randomness, for certain initial conditions: for example, if the object is cast from a roughly upright position with enough energy. A skilled thrower, they say, could have a good chance of manipulating these conditions to bias the probable outcome.

ALAMY

ASTRONOMY

When Triton lost its mate

Astron. J. **136,** 1463–1476 (2008)

Astronomers think that Neptune's largest moon, Triton, once travelled around the Sun accompanied by a planetesimal partner. At some time in the formation of the Solar System, Neptune's gravity captured Triton and ejected its mate. A new analysis details when and how this might have happened.

David Vokrouhlický of Charles University in Prague and his colleagues used the 'Nice' model, which describes how the planets attained their final configuration. According to their simulations, Triton's capture could only have occurred within the first 5–10 million years of the Solar System's formation, when a gas would have been present to slow the relative velocities of the three bodies.

That means that Neptune must have formed much earlier than some astronomers have suggested, the authors say.

ECOLOGY

Fire prevention

Nature Geosci. doi:10.1038/ngeo313 (2008)

Forest fires, whether natural or started by humans, declined worldwide during most of the past 2,000 years in parallel with

RESEARCH HIGHLIGHTS

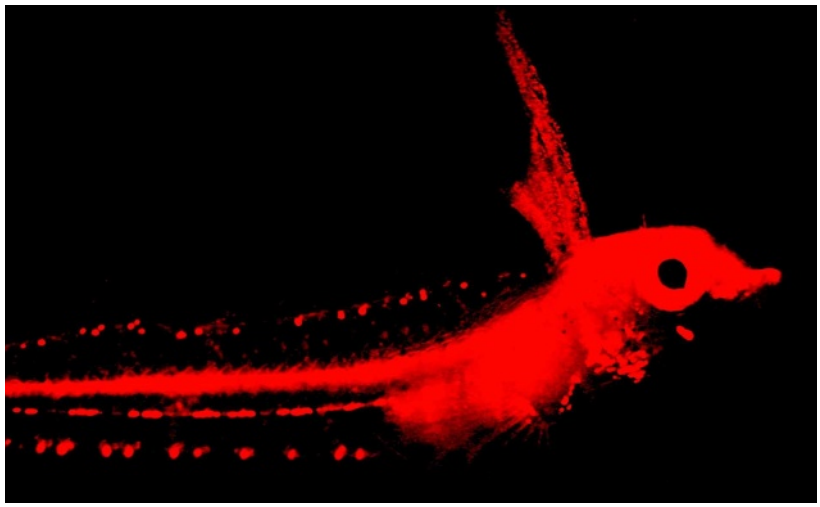
Seeing red

BMC Ecol. **8,** 16 (2008)

Water preferentially absorbs red light, so at sea it's not possible to make out red objects below a depth of around 10 metres. It was therefore thought that red was irrelevant in signalling between coral-reef fishes.

Not so. Scuba diving with filters on their masks to make red easier to pick out, Nico Michiels of the University of Tübingen in Germany and his colleagues have found at least 32 species of reef fishes that look red at depth.

These fish have crystals or pigments in their skin that fluoresce red under the incoming, mainly blue-green, light (as in the goby *Eviota destai*, pictured). The researchers say that the red markings, which are often mixed with other colours so that they appear pink, lilac or reddish brown, are most probably used for communication within each species. They add weight to this idea by showing that at least one fish, the goby *E. pellucida*, can see red.



N. K. MICHIELS *ET AL.*

CELL BIOLOGY

Home-grown fat control

Cell **134,** 933–944 (2008)

A little fat might be good for you — if it's the fatty acid palmitoleate, now discovered to be a fat-cell-derived hormone that regulates metabolism throughout the body.

Gökhan Hotamisligil at Harvard School of Public Health in Boston, Massachusetts, and his colleagues studied knockout mice lacking two proteins that normally chaperone lipids out of the bloodstream into adipose tissue. These mice are more resistant than normal mice to the bad effects of a high-fat diet. The team found increased synthesis of palmitoleate and other unsaturated fatty acids in fat cells in these mice and a consequent rise in palmitoleate levels in the blood.

They showed that palmitoleate regulates the lipid-generating program of gene expression in the liver of the mice and also stimulates muscle cells to take up glucose — and that mice infused with palmitoleate use glucose more efficiently than untreated mice.

MATERIALS SCIENCE

A rarefied insulator

Adv. Mater. **20,** 1–5 (2008)

A new material based on rare-earth elements such as cerium might overcome a barrier to making smaller silicon chips. Silicon dioxide is the traditional chip insulator, but is too bulky for smaller chips. Alternative compounds with suitably high dielectric constants are too rigid, and have proved not to insulate fully.

Now Dmitry Kukuruznyak at the Max Planck Institute for Metals Research in Stuttgart, Germany, and his colleagues have constructed another potential insulator: the rare-earth aluminium–silicon apatite $RE_6(AlO_3)_5(SiO_{3.5})$. Rather than forming rigid crystals, this compound self-organizes into flexible films on silicon at temperatures below 1,030 °C.

NONLINEAR DYNAMICS

Loading the dice

Phys. Rev. E **78,** 036207 (2008)

Gamblers — and Einstein — have assumed that throwing a die gives a random result.

But does it? Jan Nagler of the Max Planck Institute for Dynamics and Self-Organization in Göttingen and Peter Richter of the University of Bremen in Germany have simplified the throw of a die to the two-dimensional case of a dumb-bell tossed onto a surface. Will it fall with one labelled end pointing to the left or to the right?

The researchers calculated the dynamics and find that these are only truly chaotic, leading to complete randomness, for certain initial conditions: for example, if the object is cast from a roughly upright position with enough energy. A skilled thrower, they say, could have a good chance of manipulating these conditions to bias the probable outcome.

ALAMY

ASTRONOMY

When Triton lost its mate

Astron. J. **136,** 1463–1476 (2008)

Astronomers think that Neptune's largest moon, Triton, once travelled around the Sun accompanied by a planetesimal partner. At some time in the formation of the Solar System, Neptune's gravity captured Triton and ejected its mate. A new analysis details when and how this might have happened.

David Vokrouhlický of Charles University in Prague and his colleagues used the 'Nice' model, which describes how the planets attained their final configuration. According to their simulations, Triton's capture could only have occurred within the first 5–10 million years of the Solar System's formation, when a gas would have been present to slow the relative velocities of the three bodies.

That means that Neptune must have formed much earlier than some astronomers have suggested, the authors say.

ECOLOGY

Fire prevention

Nature Geosci. doi:10.1038/ngeo313 (2008)

Forest fires, whether natural or started by humans, declined worldwide during most of the past 2,000 years in parallel with

global cooling. According to a new analysis, biomass burning increased only between AD1750 and 1870, when the climate started to become warmer, populations grew rapidly, land use changed as agriculture intensified, and the industrial revolution raised atmospheric levels of carbon dioxide.

Jennifer Marlon at the University of Oregon, Eugene, and her colleagues used records of charcoal in sediments from 406 lakes, bogs and small hollows across six continents to reconstruct trends in wildfires. Wildfires declined again after 1870 as a result of landscape fragmentation and of fire management in the twentieth century.

The increase in wildfire frequency that has been observed during the past three decades — attributed to current global warming — is not yet represented in the sediment record.

NEUROSCIENCE

Sweet connections

Science **321,** 1690–1692 (2008)

When a lab rat is learning that it will get a sugar treat shortly after a light flashes, the synapses within the dopamine circuitry in its brain become temporarily more efficient.

Antonello Bonci from the Ernest Gallo Clinic and Research Center at the University of California, San Francisco and his colleagues identified this cellular mechanism involved in predicting the availability of a reward. They used, among other methods, new techniques to measure rapid changes in neurotransmitter levels *in vivo* while the rats were actually being trained.

Learning to associate an environmental signal with a reward is critical for survival. Addictive drugs short-circuit this system. They similarly increase synaptic strength, but do so for very long periods.

PALAEONTOLOGY

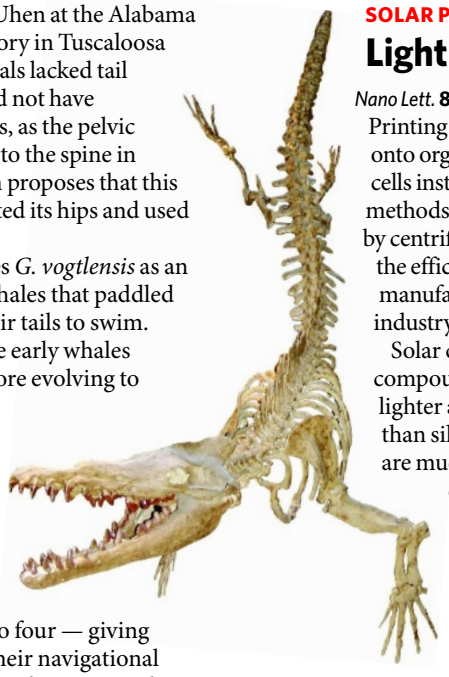
Whale of a tail

J. Vert. Paleontol. **28,** 589–593 (2008)

The ancient whale *Georgiacetus vogtlensis* probably wiggled its hips and used its giant hind feet to propel itself through the water some 40 million years ago.

Examining previously unstudied bones from this species, Mark Uhen at the Alabama Museum of Natural History in Tuscaloosa discovered that the animals lacked tail flukes. *G. vogtlensis* could not have swum by paddling its legs, as the pelvic bones are not connected to the spine in this species, and so Uhen proposes that this whale must have undulated its hips and used its feet for propulsion.

The new find identifies *G. vogtlensis* as an intermediate between whales that paddled and whales that used their tails to swim. It also suggests that some early whales undulated their hips before evolving to undulate their tails.



GEORGIA SOUTHERN MUS.

ANIMAL BEHAVIOUR

Counting bee

Anim. Cogn. **11,** 683–689 (2008)

Honeybees (*Apis mellifera*) can count up to four — giving them another string to their navigational bow. Working at the Australian National University in Canberra, Marie Dacke and Mandyam Srinivasan trained the insects to fly down a tunnel in search of food placed beside one of five identical landmarks positioned at intervals.

When trained bees flew into a tunnel that had no food, they searched most at the previously rewarding landmark — unless it was number five.

Moving the landmarks nearer to or farther away from each other did not fool the bees, showing that they were not relying on distance, but were counting the number of landmarks before the food. Changing landmarks from stripes to spots had no effect either, suggesting that bees can use numbers in an abstract way.

SOLAR POWER

Light work

Nano Lett. **8,** 2806–2813 (2008)

Printing the active coating onto organic photovoltaic solar cells instead of using other methods, such as spreading it by centrifugal force, improves the efficiency of solar-cell manufacture, according to industry scientists.

Solar cells based on organic compounds can be cheaper, lighter and more versatile than silicon-based cells, but are much less efficient at converting sunlight into electrical power.

The printing technique developed by Claudia Hoth and her colleagues at the German arm of Konarka Technologies, a solar materials manufacturer based in Lowell, Massachusetts, allows organic solar cells to be produced more simply and quickly, which might make them commercially viable.

The efficiency of cells made by this method is 3.5%, still lagging behind the 5.21% efficiency of the best organic solar cells produced by more conventional methods.

JOURNAL CLUB

Francisco Azuaje
CRP-Santé, Luxembourg

A bioinformatician considers the general applicability of host-pathogen computer simulations

Computer simulations can help explain evolutionary phenomena such as co-evolution and the emergence of robustness. Unlike traditional methods of analysis, such simulations can incorporate detailed representations of environmental antagonisms — such as the pressure that parasites exert on the evolution of their hosts.

This is what Marcel Salathé of ETH Zurich in Switzerland and Orkun Soyer of the University of Trento, Italy, recently analysed at the molecular level. By using computer simulations based on mathematical models, they showed how robust signalling networks may evolve in parasite-infested cells (M. Salathé and O. S. Soyer *Mol. Syst. Biol.* **4,** 202; 2008). In their simulations, signalling networks exhibited increasing redundancy in response to parasites, to the point that a node could be entirely removed without affecting network function. It seems that network redundancy may be a signature of parasitism present or past.

The paper is an exciting invitation to take a computational approach to evolutionary questions, by including more detailed mathematical representations. One could, for example, extend the host-parasite model to incorporate not just protein sequences, but also the ways in which genomic variation is generated, and see how everything plays out.

The approach could be generalized. National security studies, for example, might examine when and how attempts to infiltrate terrorist networks might actually make them more robust. And perhaps Salathé and Soyer's approach could be used to find ways of using environmental interference to reduce the robustness of disease networks, such as cancer signalling pathways, by examining their antagonistic interactions with therapeutic agents.

Discuss this paper at **http://blogs.nature.com/nature/journalclub**

global cooling. According to a new analysis, biomass burning increased only between AD1750 and 1870, when the climate started to become warmer, populations grew rapidly, land use changed as agriculture intensified, and the industrial revolution raised atmospheric levels of carbon dioxide.

Jennifer Marlon at the University of Oregon, Eugene, and her colleagues used records of charcoal in sediments from 406 lakes, bogs and small hollows across six continents to reconstruct trends in wildfires. Wildfires declined again after 1870 as a result of landscape fragmentation and of fire management in the twentieth century.

The increase in wildfire frequency that has been observed during the past three decades — attributed to current global warming — is not yet represented in the sediment record.

NEUROSCIENCE

Sweet connections

Science **321,** 1690–1692 (2008)

When a lab rat is learning that it will get a sugar treat shortly after a light flashes, the synapses within the dopamine circuitry in its brain become temporarily more efficient.

Antonello Bonci from the Ernest Gallo Clinic and Research Center at the University of California, San Francisco and his colleagues identified this cellular mechanism involved in predicting the availability of a reward. They used, among other methods, new techniques to measure rapid changes in neurotransmitter levels *in vivo* while the rats were actually being trained.

Learning to associate an environmental signal with a reward is critical for survival. Addictive drugs short-circuit this system. They similarly increase synaptic strength, but do so for very long periods.

PALAEONTOLOGY

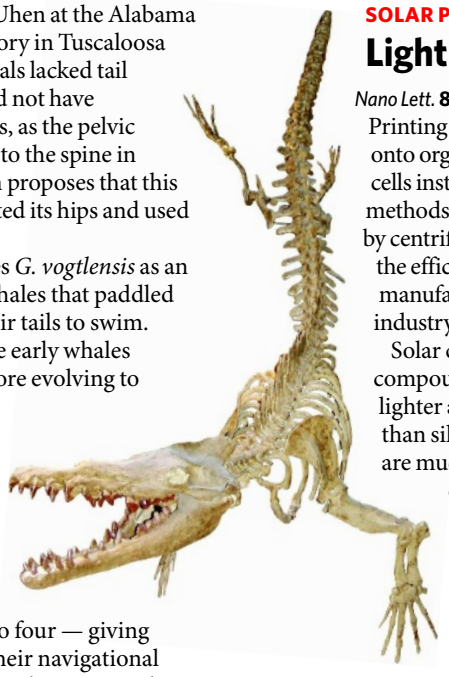
Whale of a tail

J. Vert. Paleontol. **28,** 589–593 (2008)

The ancient whale *Georgiacetus vogtlensis* probably wiggled its hips and used its giant hind feet to propel itself through the water some 40 million years ago.

Examining previously unstudied bones from this species, Mark Uhen at the Alabama Museum of Natural History in Tuscaloosa discovered that the animals lacked tail flukes. *G. vogtlensis* could not have swum by paddling its legs, as the pelvic bones are not connected to the spine in this species, and so Uhen proposes that this whale must have undulated its hips and used its feet for propulsion.

The new find identifies *G. vogtlensis* as an intermediate between whales that paddled and whales that used their tails to swim. It also suggests that some early whales undulated their hips before evolving to undulate their tails.



GEORGIA SOUTHERN MUS.

ANIMAL BEHAVIOUR

Counting bee

Anim. Cogn. **11,** 683–689 (2008)

Honeybees (*Apis mellifera*) can count up to four — giving them another string to their navigational bow. Working at the Australian National University in Canberra, Marie Dacke and Mandyam Srinivasan trained the insects to fly down a tunnel in search of food placed beside one of five identical landmarks positioned at intervals.

When trained bees flew into a tunnel that had no food, they searched most at the previously rewarding landmark — unless it was number five.

Moving the landmarks nearer to or farther away from each other did not fool the bees, showing that they were not relying on distance, but were counting the number of landmarks before the food. Changing landmarks from stripes to spots had no effect either, suggesting that bees can use numbers in an abstract way.

SOLAR POWER

Light work

Nano Lett. **8,** 2806–2813 (2008)

Printing the active coating onto organic photovoltaic solar cells instead of using other methods, such as spreading it by centrifugal force, improves the efficiency of solar-cell manufacture, according to industry scientists.

Solar cells based on organic compounds can be cheaper, lighter and more versatile than silicon-based cells, but are much less efficient at converting sunlight into electrical power.

The printing technique developed by Claudia Hoth and her colleagues at the German arm of Konarka Technologies, a solar materials manufacturer based in Lowell, Massachusetts, allows organic solar cells to be produced more simply and quickly, which might make them commercially viable.

The efficiency of cells made by this method is 3.5%, still lagging behind the 5.21% efficiency of the best organic solar cells produced by more conventional methods.

JOURNAL CLUB

Francisco Azuaje
CRP-Santé, Luxembourg

A bioinformatician considers the general applicability of host-pathogen computer simulations

Computer simulations can help explain evolutionary phenomena such as co-evolution and the emergence of robustness. Unlike traditional methods of analysis, such simulations can incorporate detailed representations of environmental antagonisms — such as the pressure that parasites exert on the evolution of their hosts.

This is what Marcel Salathé of ETH Zurich in Switzerland and Orkun Soyer of the University of Trento, Italy, recently analysed at the molecular level. By using computer simulations based on mathematical models, they showed how robust signalling networks may evolve in parasite-infested cells (M. Salathé and O. S. Soyer *Mol. Syst. Biol.* **4,** 202; 2008). In their simulations, signalling networks exhibited increasing redundancy in response to parasites, to the point that a node could be entirely removed without affecting network function. It seems that network redundancy may be a signature of parasitism present or past.

The paper is an exciting invitation to take a computational approach to evolutionary questions, by including more detailed mathematical representations. One could, for example, extend the host-parasite model to incorporate not just protein sequences, but also the ways in which genomic variation is generated, and see how everything plays out.

The approach could be generalized. National security studies, for example, might examine when and how attempts to infiltrate terrorist networks might actually make them more robust. And perhaps Salathé and Soyer's approach could be used to find ways of using environmental interference to reduce the robustness of disease networks, such as cancer signalling pathways, by examining their antagonistic interactions with therapeutic agents.

Discuss this paper at **http://blogs.nature.com/nature/journalclub**

NEWS

LHC meltdown before first collision

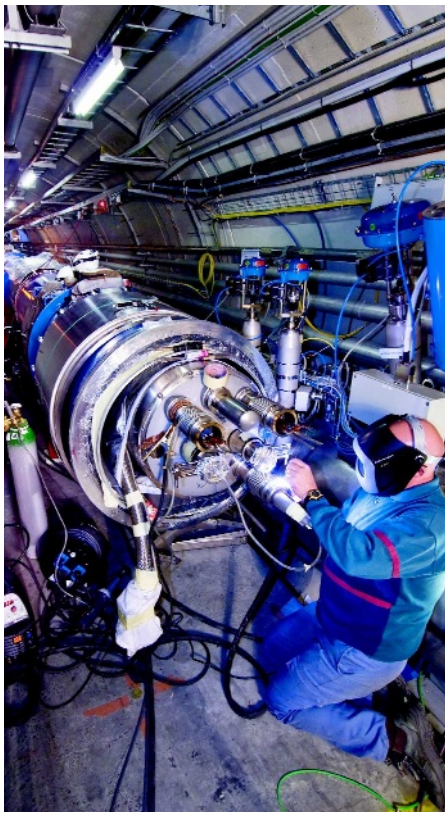
A mechanical failure has shut down the world's largest particle accelerator for at least two months.

The Large Hadron Collider (LHC), a US$4.1-billion machine located at CERN, Europe's particle-physics laboratory near Geneva, Switzerland, will be out of commission possibly until early next year to allow repairs to be made to a section of superconducting cable that is thought to have melted during a test on 19 September, just nine days after the collider circulated its first beams.

"We've had better days," says James Gillies, CERN's chief spokesperson.

The LHC is designed to accelerate protons to energies of 7 teraelectronvolts before smashing them together. To keep the beam focused and fast, the machine depends on around 9,600 superconducting niobium–titanium magnets. The magnets must be kept at temperatures of just 1.9 kelvin while carrying roughly 10,000 amps. The magnets carry this enormous load without generating heat, but if a section of the niobium–titanium rises above its operating temperature it can lead to a catastrophic failure known as a 'quench'.

That is exactly what seems to have happened as physicists passed 8,000 amps into a sector of the LHC's 27-kilometre underground ring. A cable feeding current between two of the LHC's beam-focusing quadrupole magnets suddenly heated to above superconducting temperatures and melted. The failure seems to have happened at a joint where two sections of cable were spliced together. Tens of thousands of joints run around the LHC and many of them had already been tested without incident.



M. BRICE/CERN

The shutdown may be due to a welding failure.

The failure caused the liquid helium that was being used to cool the magnets to boil off, apparently rupturing the machine and releasing as much as a tonne of the gas into the LHC tunnels. During testing the tunnels are evacuated and no injuries were reported.

Such failures are not uncommon during the early commissioning of an accelerator, Gillies says. "With a normally conducting machine you could fix it in a couple of days." But the LHC's superconducting status also makes it difficult to service. To fix the broken sector, physicists must heat thousands of tonnes of magnets from near-absolute-zero to room temperature, make the necessary repairs, and then slowly cool the system back down. Just warming and cooling will take at least two months, Gillies says.

The setback has dashed the hopes of thousands of researchers working on the four main detectors that will use the LHC. They had hoped to get their first data from proton–proton collisions as early as this week, but it now

California institute to help stem-cell biotechs

The California Institute for Regenerative Medicine (CIRM) is launching a loan programme that aims to help private biotechnology companies while providing a funding stream for the institute. The programme is expected to be approved this week by the board that oversees CIRM, although major specifics are still undecided, including the amount of money involved.

The loans are intended to help companies bring stem-cell discoveries to the clinic. It should also foster a closer relationship between California's biotech industry and the institute, which was created in 2004. Although CIRM has approved more than $614 million in grants since its inception, only one of those has gone to a company.

CIRM's could prove the biggest US state-funded biotechnology loan programme. Global accounting firm PriceWaterhouseCoopers has calculated that the institute stands to make at least $100 million over an 8- to 10-year programme, assuming that it issues loans totalling $500 million to $700 million. At a meeting on 20 September, a subcommittee of the institute's board discussed a $500-million figure for the programme. That would nicely stretch CIRM's reserves over the $3-billion, 10-year lifespan approved by voters.

The loans are intended to help companies cross regions of the "valley of death" that stops many promising ideas from making it to the clinic, says Duane Roth, chief executive of the San Diego-based business-development organization CONNECT, who headed the loan-policy task force.

"It's for after the research grant, where you have really exciting information but a lot of work to do to make it to the point where you can get traditional funders, such as venture capital and venture debt to invest," Roth says.

Few companies are trialling therapies made from stem cells yet. One, StemCells in Palo Alto, California, is testing neural stem cells derived from fetal tissue in children with the fatal genetic condition Batten disease (see *Nature* doi:10.1038/news060703-3; 2006). The company enrolled six children in the trial, but one of them died in January from the illness, according to the company. Geron in Menlo Park, California, has plans to test cells derived from embryonic stem cells in patients with spinal-cord injury, but in May the US Food and Drug Administration delayed the trial.

That raises the question of how comfortable the public will be with committing CIRM funds to one of the riskiest fields of a risky sector. Two types of loan will be awarded — recourse loans,

seems unlikely that collisions will occur this year. CERN normally shuts down during the winter months to save on electricity costs, says Tommaso Dorigo, a physicist at the CERN's Compact Muon Solenoid detector. He and others think the repairs will run into that winter shutdown. "I'm not sure collisions are going to happen in 2008," Dorigo says.

That doesn't mean physicists will be sitting on their hands in the coming months. Much must still be done to prepare the detectors for first collisions, according to Pauline Gagnon, a senior scientist on CERN's ATLAS detector. Physicists must still calibrate the hundreds of thousands of sub-detectors in each of the larger experiments, and that can be done with high-energy cosmic rays from outer space rather than beam from the machine. "It would be easier with the beam, but we can still do it," Gagnon says.

But for graduate students trying to complete their theses, there is no substitution for the real thing. Fresh data are imperative, especially for the American contingent, according to Seth Zenz, a graduate student at the University of California, Berkeley, who is currently working on ATLAS. Without new data, Zenz explains, he can't complete his thesis, and he is now planning on spending seven years in graduate school instead of six. Still, Zenz says, he's willing to put up with the delays: "I can wait a few more months before getting back to affordable sushi and burritos in California," he says. ■

Geoff Brumfiel

which must be paid back, and non-recourse loans, which do not have to be paid back if a product's development ends. CIRM is guaranteed a stake in the recipient company worth 10% of the loan's value for a recourse loan, or 100% for a non-recourse loan. The loans come with other strings — for instance, a requirement to provide California with any developed product at a state-approved price.

Greg Bonfiglio, managing partner of Palo Alto-based Proteus Venture Partners, which focuses on regenerative medicine, estimates that roughly 700 companies around the world are working in the field, and many are in California, where the loan funds must be spent. So there is already a market for the loans, he says.

What will determine biotech's response — and CIRM's potential windfall — will be their assessment of the trade-offs involved in accepting CIRM money: "It's really about the terms — that's where the rubber meets the road," Bonfiglio says. ■

Erika Check Hayden

M. SMITH/GETTY

FDA TO REGULATE GENETIC ENGINEERING
Drug agency opens door to market approval for transgenic animals.
www.nature.com/news

Monoclonal antibodies come of age

A fast way of isolating antibodies from people has been used to create a library of the immune proteins produced by someone inoculated with a nicotine-acting vaccine. Roger Beerli and his team at Cytos Biotechnology in Schlieren, Switzerland, used lymphocytes from an individual who was enrolled in a clinical trial of the smoking-cessation vaccine, and with their technique rapidly identified nicotine-specific antibodies[1].

The work is the latest offering in a burgeoning field of therapeutics: monoclonal antibodies. These antibodies are derived from a single population of cells and bind to their target at a specific site. Last month, researchers reported that they had isolated functional antibodies from survivors of the 1918 influenza pandemic[2], and in April, another team reported the rapid cloning of influenza antibodies from people who had recently been vaccinated against the disease[3]. Researchers hope the findings will eventually lead to 'passive immunity' treatments.

The market for monoclonal antibodies is the fastest-growing segment of the pharmaceutical industry. In 2007, therapeutic monoclonal antibodies brought in more than US$26 billion, most of which came from treatments for cancer and autoimmune diseases. That same year, 50 companies had anticancer antibodies in clinical trials worldwide[4].

"Infectious-disease research has come full circle."

The field has come a long way since pre-antibiotic days when infected patients were injected with a serum of horse antibodies from an animal that had been exposed to the same disease. Such treatments carried a high risk of serum sickness caused by immune reactions to the horse proteins. When small-molecule antibiotics emerged on the scene, animal serum therapies were largely abandoned.

Recent advances in antibody harvesting are breathing new life into the idea of passive-immunity therapy, this time with human antibodies. "Infectious-disease research has come full circle," says Andrew Chan, senior vice-president of immunology and antibody engineering at Genentech, a biotechnology company in South San Francisco, California.

The fundamental techniques for making monoclonal antibodies were laid down years ago, but as the market for monoclonal antibodies grows, companies are modifying those techniques to overcome technical difficulties and to establish a hold in a crowded patent landscape. "Small biotechnology companies are trying to figure out new ways to work around existing intellectual property," says immunologist James Crowe of Vanderbilt University Medical Center in Nashville, Tennessee. "That gives them a high incentive to be innovative."

Isolating antibodies from immunized humans has its limits, however. "The key is that you have to start with a disease that you can immunize people against," says Crowe. Researchers cannot, for example, immunize patients against cancer or infectious diseases for which there is no approved vaccine. And, Crowe adds, "You're not going to immunize someone against the 1918 flu". However, people who were exposed to it but survived because of their own natural antibody response might have antibodies effective against future, similar strains. Other researchers have isolated antibodies from survivors of H5N1 avian flu[5].

In theory, such antibodies could be tested for use as therapies, but antibodies are much more difficult and expensive to produce than small molecules. At present, there is only one available targeted antibody therapy against an infectious disease — MedImmune's palivizumab, used to fend off respiratory syncytial virus in premature infants. Palivizumab brings in more than a billion dollars a year for the biotech company, based in Gaithersburg, Maryland, but few other infectious diseases are expected to be so lucrative.

The high specificity of an antibody is a valued asset in the clinic, but can be detrimental in the fast-changing world of infectious disease, in which viruses such as HIV can mutate out of an antibody's grasp. "We're seeing a little bit of rediscovering large molecules like monoclonal antibodies for infectious diseases," says Chan. "But it's still quite limited."

Some companies may be wary of trialling monoclonal antibody therapies after a trial in 2006 by the now-defunct German company TeGenero in which six healthy volunteers were left fighting for their lives. But the drug (TGN1412) the six took was unusual in that it activated immune cells rather than inhibiting them.

Human-produced antibodies could also be used as a research tool, says Antonio Lanzavecchia of the Institute for Research in Biomedicine in Bellinzona, Swizerland. "I really believe this approach is going to be the major pathway to new vaccines," he says. ■

Heidi Ledford

1. Beerli, R. R. *et al. Proc. Natl Acad. Sci. USA* **105,** 14336–14341 (2008).
2. Yu, X. *et al. Nature* doi:10.1038/nature07231 (2008).
3. Wrammert, J. *et al. Nature* **453,** 667–671 (2008).
4. Reichert, J. M. & Valge-Archer, V. E. *Nature Rev. Drug Discov.* **6,** 349–356 (2007).
5. Simmons, C. P. *et al. PLoS Med.* **4,** e178 (2007).

seems unlikely that collisions will occur this year. CERN normally shuts down during the winter months to save on electricity costs, says Tommaso Dorigo, a physicist at the CERN's Compact Muon Solenoid detector. He and others think the repairs will run into that winter shutdown. "I'm not sure collisions are going to happen in 2008," Dorigo says.

That doesn't mean physicists will be sitting on their hands in the coming months. Much must still be done to prepare the detectors for first collisions, according to Pauline Gagnon, a senior scientist on CERN's ATLAS detector. Physicists must still calibrate the hundreds of thousands of sub-detectors in each of the larger experiments, and that can be done with high-energy cosmic rays from outer space rather than beam from the machine. "It would be easier with the beam, but we can still do it," Gagnon says.

But for graduate students trying to complete their theses, there is no substitution for the real thing. Fresh data are imperative, especially for the American contingent, according to Seth Zenz, a graduate student at the University of California, Berkeley, who is currently working on ATLAS. Without new data, Zenz explains, he can't complete his thesis, and he is now planning on spending seven years in graduate school instead of six. Still, Zenz says, he's willing to put up with the delays: "I can wait a few more months before getting back to affordable sushi and burritos in California," he says. ■

Geoff Brumfiel

which must be paid back, and non-recourse loans, which do not have to be paid back if a product's development ends. CIRM is guaranteed a stake in the recipient company worth 10% of the loan's value for a recourse loan, or 100% for a non-recourse loan. The loans come with other strings — for instance, a requirement to provide California with any developed product at a state-approved price.

Greg Bonfiglio, managing partner of Palo Alto-based Proteus Venture Partners, which focuses on regenerative medicine, estimates that roughly 700 companies around the world are working in the field, and many are in California, where the loan funds must be spent. So there is already a market for the loans, he says.

What will determine biotech's response — and CIRM's potential windfall — will be their assessment of the trade-offs involved in accepting CIRM money: "It's really about the terms — that's where the rubber meets the road," Bonfiglio says. ■

Erika Check Hayden

M. SMITH/GETTY

FDA TO REGULATE GENETIC ENGINEERING
Drug agency opens door to market approval for transgenic animals.
www.nature.com/news

Monoclonal antibodies come of age

A fast way of isolating antibodies from people has been used to create a library of the immune proteins produced by someone inoculated with a nicotine-acting vaccine. Roger Beerli and his team at Cytos Biotechnology in Schlieren, Switzerland, used lymphocytes from an individual who was enrolled in a clinical trial of the smoking-cessation vaccine, and with their technique rapidly identified nicotine-specific antibodies[1].

The work is the latest offering in a burgeoning field of therapeutics: monoclonal antibodies. These antibodies are derived from a single population of cells and bind to their target at a specific site. Last month, researchers reported that they had isolated functional antibodies from survivors of the 1918 influenza pandemic[2], and in April, another team reported the rapid cloning of influenza antibodies from people who had recently been vaccinated against the disease[3]. Researchers hope the findings will eventually lead to 'passive immunity' treatments.

The market for monoclonal antibodies is the fastest-growing segment of the pharmaceutical industry. In 2007, therapeutic monoclonal antibodies brought in more than US$26 billion, most of which came from treatments for cancer and autoimmune diseases. That same year, 50 companies had anticancer antibodies in clinical trials worldwide[4].

"Infectious-disease research has come full circle."

The field has come a long way since pre-antibiotic days when infected patients were injected with a serum of horse antibodies from an animal that had been exposed to the same disease. Such treatments carried a high risk of serum sickness caused by immune reactions to the horse proteins. When small-molecule antibiotics emerged on the scene, animal serum therapies were largely abandoned.

Recent advances in antibody harvesting are breathing new life into the idea of passive-immunity therapy, this time with human antibodies. "Infectious-disease research has come full circle," says Andrew Chan, senior vice-president of immunology and antibody engineering at Genentech, a biotechnology company in South San Francisco, California.

The fundamental techniques for making monoclonal antibodies were laid down years ago, but as the market for monoclonal antibodies grows, companies are modifying those techniques to overcome technical difficulties and to establish a hold in a crowded patent landscape. "Small biotechnology companies are trying to figure out new ways to work around existing intellectual property," says immunologist James Crowe of Vanderbilt University Medical Center in Nashville, Tennessee. "That gives them a high incentive to be innovative."

Isolating antibodies from immunized humans has its limits, however. "The key is that you have to start with a disease that you can immunize people against," says Crowe. Researchers cannot, for example, immunize patients against cancer or infectious diseases for which there is no approved vaccine. And, Crowe adds, "You're not going to immunize someone against the 1918 flu". However, people who were exposed to it but survived because of their own natural antibody response might have antibodies effective against future, similar strains. Other researchers have isolated antibodies from survivors of H5N1 avian flu[5].

In theory, such antibodies could be tested for use as therapies, but antibodies are much more difficult and expensive to produce than small molecules. At present, there is only one available targeted antibody therapy against an infectious disease — MedImmune's palivizumab, used to fend off respiratory syncytial virus in premature infants. Palivizumab brings in more than a billion dollars a year for the biotech company, based in Gaithersburg, Maryland, but few other infectious diseases are expected to be so lucrative.

The high specificity of an antibody is a valued asset in the clinic, but can be detrimental in the fast-changing world of infectious disease, in which viruses such as HIV can mutate out of an antibody's grasp. "We're seeing a little bit of rediscovering large molecules like monoclonal antibodies for infectious diseases," says Chan. "But it's still quite limited."

Some companies may be wary of trialling monoclonal antibody therapies after a trial in 2006 by the now-defunct German company TeGenero in which six healthy volunteers were left fighting for their lives. But the drug (TGN1412) the six took was unusual in that it activated immune cells rather than inhibiting them.

Human-produced antibodies could also be used as a research tool, says Antonio Lanzavecchia of the Institute for Research in Biomedicine in Bellinzona, Swizerland. "I really believe this approach is going to be the major pathway to new vaccines," he says. ■

Heidi Ledford

1. Beerli, R. R. *et al. Proc. Natl Acad. Sci. USA* **105,** 14336–14341 (2008).
2. Yu, X. *et al. Nature* doi:10.1038/nature07231 (2008).
3. Wrammert, J. *et al. Nature* **453,** 667–671 (2008).
4. Reichert, J. M. & Valge-Archer, V. E. *Nature Rev. Drug Discov.* **6,** 349–356 (2007).
5. Simmons, C. P. *et al. PLoS Med.* **4,** e178 (2007).

Gulf states plan for nuclear future

While Iran grabs international headlines for its controversial nuclear programme, ambitious plans for an atomic roll-out on the other side of the Gulf are quietly under way. Non-proliferation experts are already voicing concern.

Major oil-producing states, including Saudi Arabia, the United Arab Emirates (UAE) and Bahrain, are taking steps to develop civilian nuclear power. Concurrently, the Gulf Cooperation Council (GCC), a trade bloc of the six Arab states in the region, is working with the International Atomic Energy Agency (IAEA) to erect a research reactor on the peninsula. Nearly everyone in the Gulf has declared some interest in nuclear power over the past two years, says Mustafa Alani, a senior consultant with the Gulf Research Centre, an independent think tank in Dubai, UAE. "There is a fever in the region."

Interest in nuclear power is on the rise worldwide, but the Gulf may be better situated than most to actually build reactors. Whereas Western utilities struggle to raise the billions needed for a power plant, the cash-flush Gulf states have more than enough money, says Seth Grae, chief executive of Thorium Power, a consultancy on nuclear power based in McLean, Virginia. "Right now, the Gulf region represents one of the most exciting and promising nuclear markets in the world."

Grae says that the nuclear fever is driven by the region's rapidly expanding economies and the need for fresh water, for which energy-intensive desalination is the best option. But some non-proliferation experts say that fear of Iran, which may be close to developing a nuclear weapon, is the real reason behind the growing interest. "States are not driven to build reactors solely based on economic concerns," says Jill Parillo, deputy director for security programmes at Physicians for Social Responsibility, an advocacy group based in Washington DC. "It's a perceived security assurance."

It may seem strange that a region flush with oil and natural gas — not to mention the solar-energy potential — would even consider building nuclear power stations. Nuclear power plants are enormously expensive to build — recent estimates run at US$7 billion per plant. And there are time-consuming, political complications. A nation with nuclear power must introduce legislation to address liability and safety concerns, for example, and an independent regulatory body to oversee plant operations.

Yet nuclear power makes sense for the Gulf states, says Hans-Holger Rogner, the head

Brazil braced for unexpected oil wealth

At a time when most countries are struggling with ageing oilfields, Brazil is debating a new future as an oil exporter following the recent discovery of world-class reservoirs off its coast.

President Luiz Inácio Lula da Silva (Lula) has appointed a cabinet-level committee to come up with recommendations on proposals such as the creation of a state-owned company to manage the development, and a national 'rainy-day' fund for collecting and disbursing the revenue. The committee is expected to report later this week, but Lula is said to favour both approaches, and says that he wants to see the money invested in education and social programmes.

Some experts caution that attempts to consolidate control over the new oil resources could effectively undercut Brazil's biggest asset, Petróleo Brasileiro (Petrobras), the government-controlled oil company that has pioneered exploration of the new fields. Georges Landau, who heads São Paulo-based energy consultancy Prismax Consultoria, says the government could make a good case for increasing its share of the revenue now that Petrobras has confirmed there is oil to be tapped. And Brazil should create a better system for handling offshore oil revenue, much of which is currently squandered by Rio de Janeiro state, he says.

But Landau thinks the proposal to create a state-owned company is a political ploy to ensure control, driven by the fact that many foreign investors hold shares in Petrobras. In theory, the new company would manage the development while Petrobras and others do the drilling. Landau says such a move could damage Petrobras and ultimately delay the development. "That's what I fear will happen, and it will be a disaster, really. It's self-destructive."

Others are less pessimistic. "I don't think Brazil's plan is to go down the same nationalistic route of other countries in the region" (such as Bolivia, which nationalized its energy industry and shut down foreign firms), says Roger Tissot, an expert on Latin America based in British Columbia, Canada. "I think

HAVE YOUR SAY
Comment on this or any of our News stories, online.
www.nature.com/news

G. HELLIER/GETTY IMAGES

The speed of development in Dubai makes many think its nuclear power project will also be rapid.

of planning and economic studies at the IAEA, which is responsible for both overseeing nuclear safeguards and promoting the development of nuclear power. The region faces unprecedented growth over the coming decade. The UAE alone expects its electricity needs to almost quadruple, rising from around 15 gigawatts (thousand megawatts) in 2006 to nearly 40 gigawatts in 2020. Part of that demand is driven by the UAE's rapid economic development, but demand for fresh water, which will be desalinated using energy-intensive reverse osmosis, is also a major factor. The need for electricity could be met in part by burning fossil fuels, but their anticipated future price abroad makes that option uneconomical, Rogner says. "The oil and gas the Gulf states produce are worth far more on the international market than if they burn them at home."

The UAE has progressed furthest in its plans to develop nuclear power. Earlier this year, the small, oil-rich nation penned a cooperative agreement with France to begin nuclear development, and it is looking at a bid from French companies Areva, Suez and Total to build and operate two 1,600-megawatt reactors on islands off its coast. At the same time, it is enacting legislation on nuclear issues and setting up an independent regulatory body to license and oversee its plants. "They're taking it very seriously," says Grae, whose firm consults with the UAE government on its nuclear plans.

Some are concerned about the UAE programme's rapid development. The tiny state is famous for importing foreign talent to build its roads and bridges, and it looks as though nuclear will be no exception. "They're apparently talking about just buying in regulators from other countries," says Charles Ferguson of the Council on Foreign Relations, a think tank based in Washington DC. Ferguson says that these foreign regulators would face "enormous pressure" from the government if they tried to shut down a reactor for safety reasons.

Grae counters that the UAE already has a proven record of importing international talent to operate independent institutions. Dubai's International Financial Centre, for example, is administered according to international 'common law' and has its own courts staffed by foreign judges. "The UAE has some real examples of regulatory independence," Grae says.

Like many states in the region, the UAE has signed up to the international Nuclear Non-Proliferation Treaty, a document that promises aid for civilian programmes if countries give up their right to nuclear arms. In addition, it and many other GCC states have announced that they will not develop the nuclear technologies of most concern: uranium enrichment and spent-fuel reprocessing. Both technologies can be used to manufacture reactor fuel, but they can also be used to refine weapons-grade nuclear material. Rogner says he is satisfied that the Gulf states are not a danger. "A nuclear power plant per se is not a proliferation risk," he says.

Grae says that he believes economics is the sole driver behind the Gulf's push for nuclear power. "What I am seeing in the UAE is that this has zero to do with security," he says.

But Parillo has her doubts. Creating the technological basis for a civilian programme will also be of some use should the states ever need to develop nuclear weapons, she says. Building and operating plants will help nations develop expertise in areas such as reactor physics and the handling of nuclear materials. "This is the first step on the way to developing a dual-use capability."

Alani adds that it is no coincidence that the Gulf's nuclear renaissance began in earnest two years ago. In the spring of 2006, the IAEA announced that Iran was enriching uranium without the agency's approval. The Gulf states view Iran as an expansionist and interventionist regime, he says, and the international community's failure to control its programme set off alarm bells. The message to many states was clear, he says: "One has to start a civilian programme to start narrowing the gap." ■

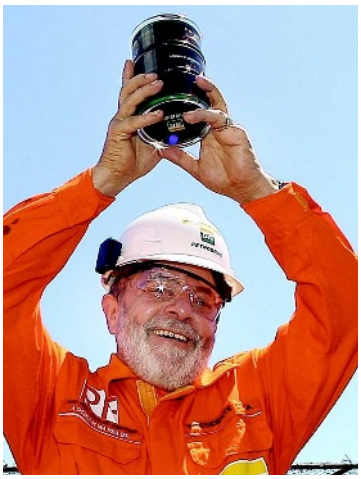
Geoff Brumfiel

Brazil is being very professional and very serious in its approach."

Tissot points out that Brazil is well within its rights to rethink the structure of its industry going forward, and he credits the government with looking to Norway as a potential model. Norway has a national fund to manage its oil revenue in perpetuity, he says, and at one time had a state-owned company — separate from Statoil, the main Norwegian oil producer — managing the development of its North Sea resources.

Petrobras controls some 95% of the nation's oil production, which is largely offshore. The new discoveries, some of which lie directly below current reservoirs south of Rio de Janeiro, are located several thousand metres below sea level in more than 2,000 metres of water. They have been dubbed 'pre-salt' fields because they also lie below a 2,000-metre layer of salt.

It's not yet clear just how much oil can be recovered, in part because nobody has ever drilled in such conditions. But some experts are setting the potential reserves at 80 billion barrels or more. If true, this would take Brazil past Russia and into the league of some of the Middle Eastern oil superpowers. Petrobras has confirmed a 5-billion to 8-billion-barrel estimate for one of the fields, but has not released official estimates for many others.



R. STUCKERT/AFP/GETTY IMAGES

President Lula brandishes a first sample from Brazil's latest oilfield.

Matthew Shaw, a Latin American analyst for Wood Mackenzie, a consultancy firm in Edinburgh, UK, says that Petrobras and its partners have drilled two wells into one field and one into each of seven other fields; ExxonMobil is getting ready to drill its own wells into an adjacent area.

"I've seen estimates of 50 billion barrels, and I've seen estimates of 80 billion barrels, but nobody knows," says Shaw. "It's definitely in the tens of billions of barrels, but there's no point in guessing now." ■

Jeff Tollefson

SNAPSHOT

Marine marvels

Neil Bruce of the Museum of Tropical Queensland in Townsville, Australia, peers into a lit aquarium on Lizard Island Reef, part of the Great Barrier Reef off the east coast of Australia. The research trip, which also went to Ningaloo reef off the west coast, turned up hundreds of new species, including dozens of new crustaceans.

The expedition forms part of the Census of Marine Life. It included the first systematic survey of soft corals, discovering more than 100 that were previously unknown. From top to bottom: Comb jelly (*Ctenophore*); green-banded snapping shrimp (*Alpheus parvirostris*); twisted nudibranch (*Chromodoris elizabethina*); *Dendronephthya* soft coral. ■

ALL IMAGES: G. CRANITCH/QUEENSLAND MUSEUM

Creationism row forces out UK educator

The director of education at Britain's Royal Society has been forced to resign after a massive outcry in the wake of widespread misreporting of comments he made about creationism in the classroom.

Michael Reiss, a professor at London's Institute of Education and an ordained minister in the Church of England, made the remarks at the British Association for the Advancement of Science's annual Festival of Science on 11 September in Liverpool. Three Nobel-prizewinning society fellows wrote to the society's president, Martin Rees, saying that they were "greatly concerned" at media reports of Reiss's talk. They are Richard Roberts, chief scientific officer of New England BioLabs in Beverly, Massachusetts; Harold Kroto of Florida State University in Tallahassee; and John Sulston of the University of Manchester, UK.

The Royal Society initially insisted that Reiss had been misrepresented and that his views do not differ from the society's position that "creationism has no scientific basis and should not be part of the science curriculum". After the letter of complaint and with the reported statements continuing to receive press coverage, including hostile opinion pieces, the society announced Reiss's departure on 16 September.

See Editorial, page 431, and http://tinyurl.com/4zluv9 for a longer version of this story.

Arctic ice shrinks less this year than last

AP PHOTO/THE CANADIAN PRESS/J. HAYWARD

Sea-ice cover in the Arctic (pictured) has reached its annual low — and not broken last year's record of the smallest ice extent since satellite records began, says the US National Snow and Ice Data Center in Boulder, Colorado.

Ice cover began to grow again from 10 to 14 September after bottoming out at 4.52 million square kilometres, the centre says. That's 9.4% more ice than last summer's minimum. Contributing factors include the fact that there were fewer warm days in the Arctic region this year than last, and also that winds blew in different directions instead of packing the ice together into a small area.

The International Arctic Research Center in Fairbanks, Alaska, also recorded a minimum last week: 4.71 million square kilometres on 9 September.

Anthrax enquiry to undergo independent review

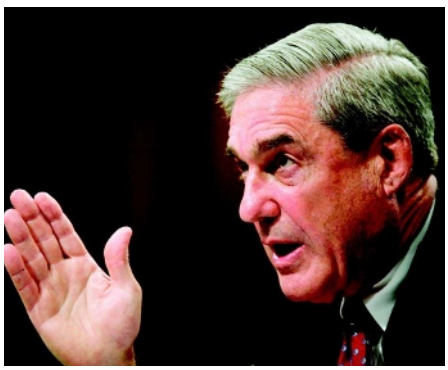
On the request of the Federal Bureau of Investigation, experts convened by the US National Academy of Sciences will review the FBI's scientific evidence against bioweapons researcher Bruce Ivins, who committed suicide earlier this year, in the anthrax bioterrorist attacks of 2001.

The scope and timeline have not yet been defined, but the assessment will exclude the 60 or so experts who consulted with the FBI on the case. Bureau director Robert Mueller revealed the planned evaluation in testimony before judiciary committees of the Senate and House of Representatives last week.

Several lawmakers have expressed doubts about the FBI's case, questioning how the bureau eliminated all other possible suspects, and why it took so long to zero in on Ivins, a researcher at the US Army Medical Research Institute of Infectious Diseases in Fort Detrick, Maryland. Senator Patrick Leahy (Democrat, Vermont), a target in the mailings of deadly spores, said he believes the attacks were not a one-man job.

Mueller responded that the FBI had followed every lead; the case is expected to last for several months.

REUTERS/M. RILEY



FBI director Robert Mueller.

Researcher punished for misconduct wins NSF grant

Days after being sanctioned for research misconduct, bubble-fusion researcher Rusi Taleyarkhan was back in business — with a $185,000 grant from the US National Science Foundation (NSF).

According to information on the foundation's website, Taleyarkhan, who is at Purdue University in West Lafayette, Indiana, has been contracted to work with students to develop a prototype particle detector based on the effect of radioactive particles on stressed fluids. Taleyarkhan's bubble-fusion experiments also involved the use of radioactive particles to seed bubbles in fluids.

In July, Taleyarkhan was found guilty of two charges of falsification and on 27 August Purdue banned him from having graduate students for three years, and withdrew his named professorship.

But the university apparently failed to notify NSF programme manager Rajinder Khosla, who confirms that Taleyarkhan's grant began on 15 September. Khosla says that he was not aware of the controversy, and plans to review the situation.

Australian clinic gets permit to clone stem cells

The *in vitro* fertilization clinic Sydney IVF has received Australia's first licence to conduct therapeutic cloning. The centre's team will be one of a handful of groups worldwide trying to achieve the feat.

In therapeutic cloning, also known as somatic-cell nuclear transfer, DNA from a patient's cell is introduced into an unfertilized egg. When the egg develops into an embryo, scientists extract stem cells and try to create an embryonic-stem-cell line that is a genetic match of the patient.

The licence provides for use of 7,200 human eggs over the next 3 years. Most other groups that have tried and failed have had only 20–30 eggs. Julia Schaft, who will lead the group, says it should be able to use at least 3,500 "clinically unsuitable" eggs from the 20,000 total the clinic collects each year.

Agency drops disputed chelation study

A controversial trial of chelation therapy for autism (see *Nature* 454, 259; 2008) has been scrapped by the US National Institute of Mental Health, which says its resources are better directed elsewhere.

The study was first proposed in 2006, but was referred for further ethics review after a 2007 publication reported that chelating agents could cause cognitive problems in rats. Critics charged that the study would needlessly expose children with autism to risk for no medical gain.

The American Academy of Pediatrics has recommended that the treatment be reserved for children with serious heavy-metal poisoning. Still, it is in broad underground use by parents of children with autism, many of whom believe the disease is due to mercury in childhood vaccinations.

Creationism row forces out UK educator

The director of education at Britain's Royal Society has been forced to resign after a massive outcry in the wake of widespread misreporting of comments he made about creationism in the classroom.

Michael Reiss, a professor at London's Institute of Education and an ordained minister in the Church of England, made the remarks at the British Association for the Advancement of Science's annual Festival of Science on 11 September in Liverpool. Three Nobel-prizewinning society fellows wrote to the society's president, Martin Rees, saying that they were "greatly concerned" at media reports of Reiss's talk. They are Richard Roberts, chief scientific officer of New England BioLabs in Beverly, Massachusetts; Harold Kroto of Florida State University in Tallahassee; and John Sulston of the University of Manchester, UK.

The Royal Society initially insisted that Reiss had been misrepresented and that his views do not differ from the society's position that "creationism has no scientific basis and should not be part of the science curriculum". After the letter of complaint and with the reported statements continuing to receive press coverage, including hostile opinion pieces, the society announced Reiss's departure on 16 September.

See Editorial, page 431, and http://tinyurl.com/4zluv9 for a longer version of this story.

Arctic ice shrinks less this year than last

AP PHOTO/THE CANADIAN PRESS/J. HAYWARD

Sea-ice cover in the Arctic (pictured) has reached its annual low — and not broken last year's record of the smallest ice extent since satellite records began, says the US National Snow and Ice Data Center in Boulder, Colorado.

Ice cover began to grow again from 10 to 14 September after bottoming out at 4.52 million square kilometres, the centre says. That's 9.4% more ice than last summer's minimum. Contributing factors include the fact that there were fewer warm days in the Arctic region this year than last, and also that winds blew in different directions instead of packing the ice together into a small area.

The International Arctic Research Center in Fairbanks, Alaska, also recorded a minimum last week: 4.71 million square kilometres on 9 September.

Anthrax enquiry to undergo independent review

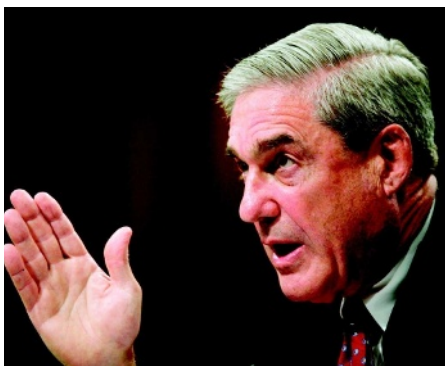
On the request of the Federal Bureau of Investigation, experts convened by the US National Academy of Sciences will review the FBI's scientific evidence against bioweapons researcher Bruce Ivins, who committed suicide earlier this year, in the anthrax bioterrorist attacks of 2001.

The scope and timeline have not yet been defined, but the assessment will exclude the 60 or so experts who consulted with the FBI on the case. Bureau director Robert Mueller revealed the planned evaluation in testimony before judiciary committees of the Senate and House of Representatives last week.

Several lawmakers have expressed doubts about the FBI's case, questioning how the bureau eliminated all other possible suspects, and why it took so long to zero in on Ivins, a researcher at the US Army Medical Research Institute of Infectious Diseases in Fort Detrick, Maryland. Senator Patrick Leahy (Democrat, Vermont), a target in the mailings of deadly spores, said he believes the attacks were not a one-man job.

Mueller responded that the FBI had followed every lead; the case is expected to last for several months.

REUTERS/M. RILEY



FBI director Robert Mueller.

Researcher punished for misconduct wins NSF grant

Days after being sanctioned for research misconduct, bubble-fusion researcher Rusi Taleyarkhan was back in business — with a $185,000 grant from the US National Science Foundation (NSF).

According to information on the foundation's website, Taleyarkhan, who is at Purdue University in West Lafayette, Indiana, has been contracted to work with students to develop a prototype particle detector based on the effect of radioactive particles on stressed fluids. Taleyarkhan's bubble-fusion experiments also involved the use of radioactive particles to seed bubbles in fluids.

In July, Taleyarkhan was found guilty of two charges of falsification and on 27 August Purdue banned him from having graduate students for three years, and withdrew his named professorship.

But the university apparently failed to notify NSF programme manager Rajinder Khosla, who confirms that Taleyarkhan's grant began on 15 September. Khosla says that he was not aware of the controversy, and plans to review the situation.

Australian clinic gets permit to clone stem cells

The *in vitro* fertilization clinic Sydney IVF has received Australia's first licence to conduct therapeutic cloning. The centre's team will be one of a handful of groups worldwide trying to achieve the feat.

In therapeutic cloning, also known as somatic-cell nuclear transfer, DNA from a patient's cell is introduced into an unfertilized egg. When the egg develops into an embryo, scientists extract stem cells and try to create an embryonic-stem-cell line that is a genetic match of the patient.

The licence provides for use of 7,200 human eggs over the next 3 years. Most other groups that have tried and failed have had only 20–30 eggs. Julia Schaft, who will lead the group, says it should be able to use at least 3,500 "clinically unsuitable" eggs from the 20,000 total the clinic collects each year.

Agency drops disputed chelation study

A controversial trial of chelation therapy for autism (see *Nature* 454, 259; 2008) has been scrapped by the US National Institute of Mental Health, which says its resources are better directed elsewhere.

The study was first proposed in 2006, but was referred for further ethics review after a 2007 publication reported that chelating agents could cause cognitive problems in rats. Critics charged that the study would needlessly expose children with autism to risk for no medical gain.

The American Academy of Pediatrics has recommended that the treatment be reserved for children with serious heavy-metal poisoning. Still, it is in broad underground use by parents of children with autism, many of whom believe the disease is due to mercury in childhood vaccinations.

The home stretch

The leading US presidential candidates are not trying to woo voters with science issues. But the senator who wins will help shape the world's most influential research agenda. **Alexandra Witze** looks at how John McCain and Barack Obama have developed their thoughts on science and technology, and where each of them might take the country if elected.

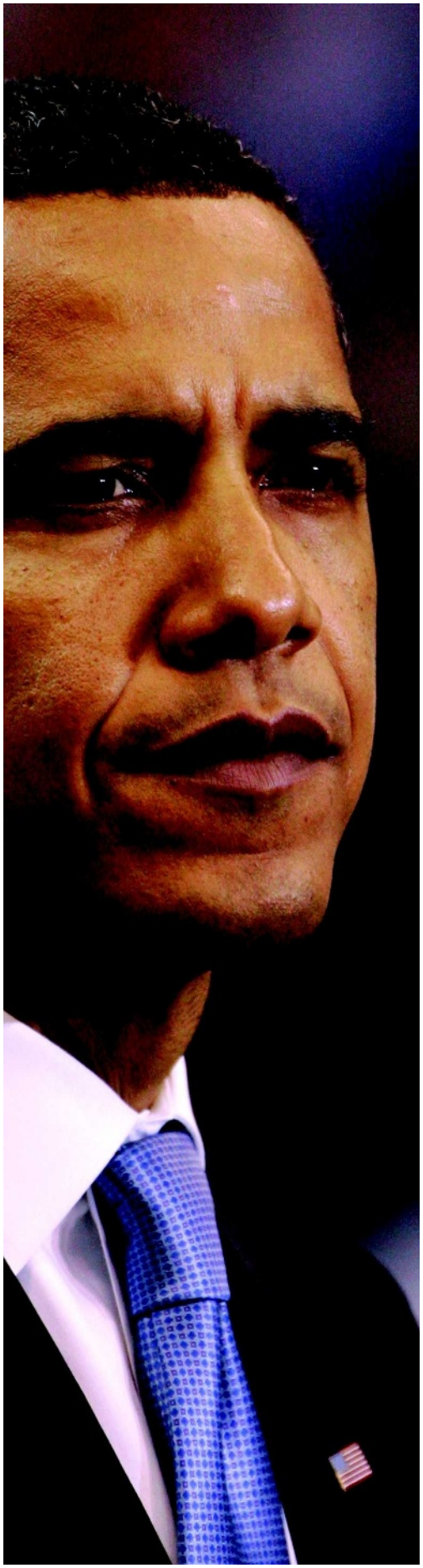
The United States' scientific establishment has been doing its best to get this year's presidential contenders to pay attention to it. In April, a group of supporters calling itself Science Debate 2008 reserved a large hall at the Franklin Institute in Philadelphia, hoping the whole field of candidates would show up to take questions. None accepted. In fact, it took evangelical pastor Rick Warren to get the two leading candidates — Republican John McCain and Democrat Barack Obama — together for the first time, at a forum at a California church in August.

It is no surprise that the campaign has revolved more around the war and the economy, and of course personality, than science. It would be a mistake, however, to assume that neither McCain nor Obama cares about science and technology. In fact, both have a long history of engaging with core science issues — from McCain running a Senate committee that oversaw science, among many topics, to Obama working on technological approaches to fighting poverty. During the course of the campaign both candidates have laid out where they stand on scientific and technical issues (see page 446).

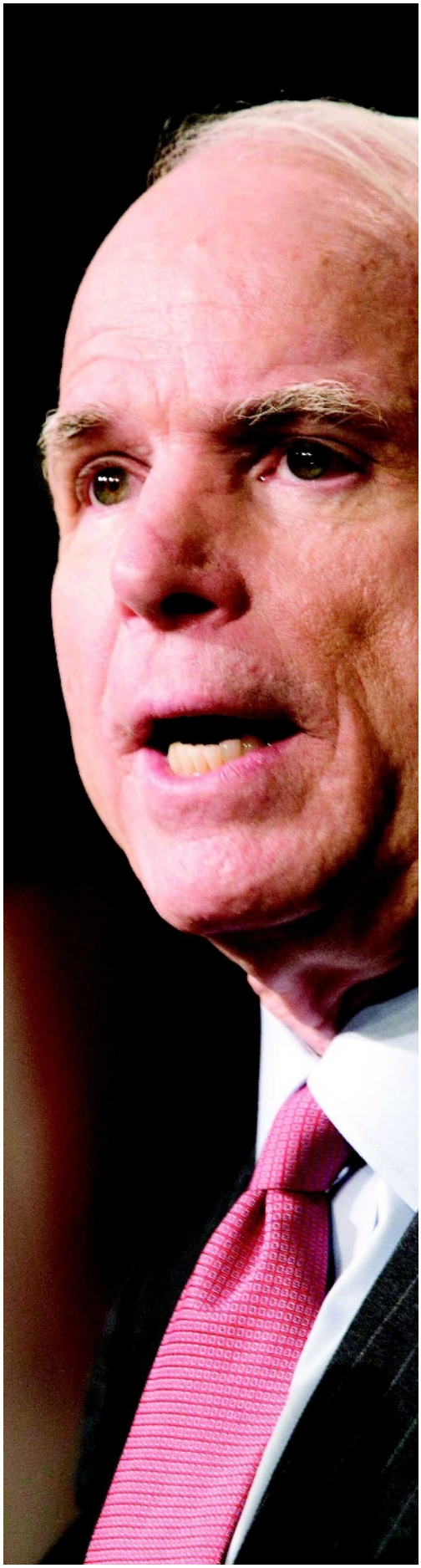
Major changes are in store no matter who wins. Many scientists argue that the research environment will be better off during the next four years than it has been under President George W. Bush's two terms. The next president will probably play a leading part in initiating the first mandatory greenhouse-gas regulations in the United States. He may also lift the restrictions on research with human embryonic stem cells that limit federal funding to cell lines that are at least seven years old. New people will move into critical positions such as the presidential science adviser (see page 453) and heads of the various science-related agencies (see page 451). This will be a breath of fresh air to researchers who charge that the Bush administration has manipulated science to political ends, from climate-change reports to endangered-species listings.

Bush aside, though, the question remains as to how McCain or Obama might influence research. What the candidates say about science will normally be broad-brush; no one expects a president to rattle off details of the confinement of quarks or the dynamics of proteins. More pertinent is how they think about science as a process, where they get their scientific advice from — and how they might incorporate such advice into where they would lead the country.

Obama's views on how science and technology affect the nation's growth trace back several years. Working with disadvantaged children



For our election podcasts and more online content, see www.nature.com/uselection.



M. SPENCER GREEN/AP; NEWSCOM/PHOTOSHOT/GARRISON PHOTOGRAPHY

on Chicago's South Side, "he saw the role and importance of technology in the future of those schoolchildren", says Alec Ross, a technology entrepreneur and campaign adviser. Early in his campaign, Obama released details on how he would expand broadband access as a way to shrink the 'digital divide' between rich and poor, and described plans to appoint the first-ever national 'chief technology officer' to improve infrastructure.

The theme of how technology investment can drive economic growth comes up again and again in Obama's speeches. His advisers say that he spends a lot of time thinking about how science and technology can be used to address specific national aims, such as health care and climate change. To this end, Obama voted in the Senate in favour of the 2007 America COMPETES Act, which authorized among other things a doubling in the physical-sciences research budgets at the National Science Foundation, the Department of Energy's Office of Science and the National Institute of Standards and Technology. It was a politically popular act that sailed through the Senate — also garnering McCain's support — perhaps because it was simply an expression of support for funding in these areas and not an actual allocation of dollars.

Expert advice

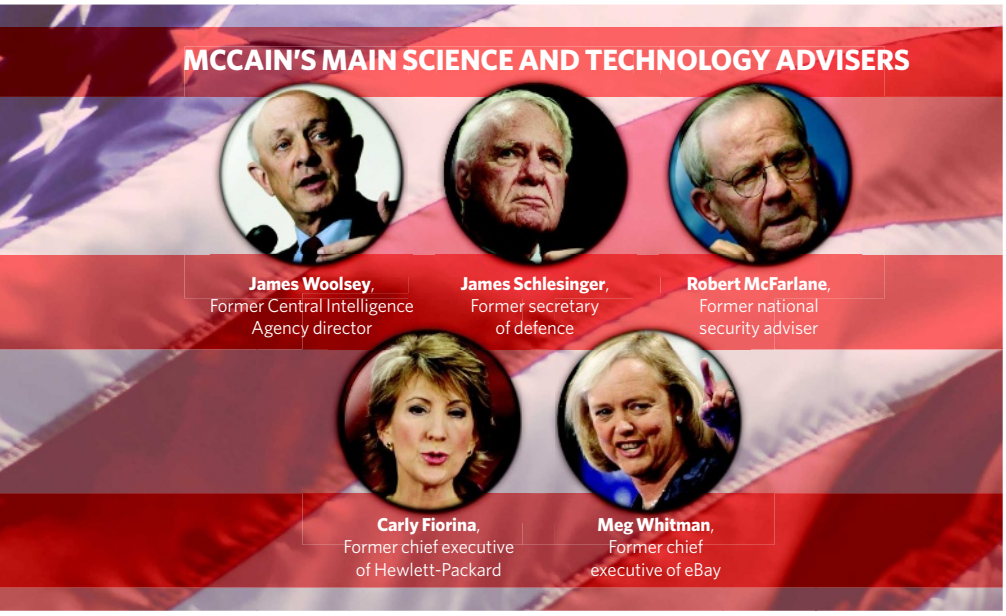
Much of Obama's advice on economic matters comes from a group of advisers that includes Austan Goolsbee of the University of Chicago in Illinois. Obama also maintains a select group of science advisers headed by Nobel laureate Harold Varmus, head of the Memorial Sloan-Kettering Cancer Center in New York and former director of the National Institutes of Health. Varmus endorsed Obama back in February and has been his linchpin for science advice since; even so, the most face time Varmus has had with the candidate was a two-hour summit on competitiveness issues at Carnegie Mellon University, in Pittsburgh, Pennsylvania, in June.

Varmus has strong local support in Don Lamb, an astrophysicist at the University of Chicago who has known Obama since he represented Lamb in the state senate of Illinois. Lamb serves as a sort of secretary for the advisory group, maintaining e-mail lists and helping coordinate discussions and the flow of advice back into the Obama campaign. Other members of the team include Henry Kelly, president of the Federation of American Scientists in Washington DC, who served as assistant director for technology in Bill Clinton's advisory office; Gil Omenn, a professor of internal medicine and human genetics at the University of Michigan in Ann Arbor who held several key policy positions in the Carter administration; and Sharon Long, a biologist at Stanford University in Palo Alto, California. Also loosely connected are a host of Nobel laureates including Bob Horvitz of the Massachusetts Institute of Technology; Peter Agre of the Johns Hopkins Malaria Research Institute in Baltimore, Maryland; and Burton Richter of the Stanford Linear Accelerator Center in Menlo Park, California. The team also includes Tom Kalil, special assistant to the chancellor for science and technology at the University of California, Berkeley, and deputy assistant to President Clinton for technology and economic policy. Kalil was the main architect of Hillary Clinton's science platform, generally regarded as the most well-developed set of policies on science for any of the presidential candidates until she dropped out of the race.

Varmus's team helps to write material that serves as the basis for developing policy platforms, and otherwise throws general science advice into the churning maw that is any candidate's advisory system. A related group funnels climate and energy advice up the same chain. This group includes energy researcher Dan Kammen of the University of California, Berkeley, and energy policy expert Jason Grumet, founder and president of the Bipartisan Policy Center in Washington DC.

Academics are notoriously left-leaning — a national study done last year by sociologists Neil Gross, from Harvard University in Cambridge, Massachusetts, and Solon Simmons, from George Mason University in Fairfax, Virginia, found that half of them identified themselves as Democrats, and just 14% as Republicans. So perhaps it's not surprising that McCain's science and technical advice comes more from corporate and political leaders than from the academic world. Much of his campaign's day-to-day work on technical matters comes from experienced staff members who have worked with McCain for years. These include Floyd DesChamps of the Senate Committee on Commerce, Science and Transportation, which McCain headed for much of the time between 1997 and 2001. The committee has a broad remit but includes, for instance, NASA, where McCain was involved in following up on the 2003 *Columbia* space-shuttle disaster. Other technical advice to the McCain campaign comes from business leaders such as Carly Fiorina, the former chief executive of Hewlett-Packard, and Meg Whitman, the former chief executive of eBay.

McCain's most detailed positions on science issues come in the field of energy. According to James Woolsey, director of central intelligence under Bill Clinton, the campaign's top policy advisers request specific advice on energy from a small group that includes Woolsey; Ronald Reagan's national security adviser Robert McFarlane (now a renewable-energy



Vice-presidential nominees: Republican Sarah Palin and Democrat Joe Biden.

LONG: L. A. CICERO/STANFORD NEWS SERVICE; KAMMEN: UC BERKELEY; WOOLSEY: AP PHOTO; SCHLESINGER: S. JAFFE/AFP/GETTY IMAGES; MCFARLANE: AP PHOTO/KEYSTONE, M. TREZZINI; FIORINA: R. EDMONDS/AP; WHITMAN: C. NEIBERGALL/AP

advocate); and James Schlesinger, secretary of defence under Richard Nixon and Gerald Ford and secretary of energy under Jimmy Carter. Since July, the McCain campaign has also relied on chief strategist Steve Schmidt — an adviser to former White House deputy chief of staff Karl Rove and to Vice-President Dick Cheney — who helped develop energy and climate policies when he masterminded the successful 2006 re-election campaign of California governor Arnold Schwarzenegger.

Environmental groups have not been big fans of McCain in recent years — yet in some respects, McCain has adopted politically risky energy policies. He has proposed scrapping subsidies and tariffs for biofuels, an approach favoured by Obama that plays well with voters in America's corn-growing heartland, but that a growing consensus among energy and environmental experts rejects. Most strikingly, McCain is a long-time supporter of a cap-and-trade approach to regulating greenhouse-gas emissions, in which the government sets an overall cap on pollution and then issues permits to each polluter; those that reduce their emissions cheaply can sell the extra permits to anybody who needs them. Whereas McCain's plan for such a system would distribute the allowances and allow businesses to sell any unused emissions for cash, Obama's version of the scheme would auction off the allowances from the start. Obama has also pledged $150 billion over the next ten years towards energy technologies.

Environmental back-up

Five years ago, McCain introduced the first version of a significant emissions-regulating bill with Joseph Lieberman, the formerly Democratic, now Independent, senator from Connecticut. Introduced three times into the Senate, the McCain–Lieberman bill never garnered enough votes to pass, but is seen as the prototype and inspiration for much of what followed. Tim Profeta, the former Lieberman staff member who worked with McCain aides to write the bill, says he thinks McCain will continue to fight for a cap-and-trade strategy if elected as president. "He may have to make some deals, but I don't see any indication from him that he really is wavering in his dedication to the topic. At this point, I have to take him at his word," says Profeta, now director of the Nicholas School of the Environment at Duke University in Durham, North Carolina.

If the presidential candidates' views on scientific matters are of only minor interest to voters, those of the vice-presidential candidates are usually even less so. This year, though, that might not hold true. Obama's choice of Joe Biden had much more to do with Biden's experience in foreign policy than anything he might have to say about offshore oil drilling. Biden, a senator from Delaware, hews to traditional Democratic lines on most technical issues. Despite his personal belief that human lives begin at conception, Biden supports abortion rights and human embryonic stem-cell research. Biden has never been shy about speaking his mind; in an appearance last month on

REUTERS/J. GRESS

REUTERS/B. SNYDER

Bill Maher's liberal television talk show, he spoke out against teaching intelligent design in science class, saying "I refuse to believe the majority of people believe this malarkey."

McCain's choice of Sarah Palin as a vice-president is much more challenging to parse. Far more than McCain himself, Palin is truly representative of the evangelical Christian right. Also convinced in the embryo's full personhood, she opposes human embryonic stem-cell research, as well as abortion for any reason other than saving the mother's life. Biden even got into a dust-up with Palin this month by asking why she doesn't support stem-cell research when she has a child with Down syndrome.

When running for governor of Alaska in 2006, Palin said that creationism should be taught in school on the basis that "healthy debate is so important". And in her only major interview to date, Palin engaged in a back-and-forth with ABC's television news anchor Charlie Gibson over whether she had ever challenged humanity's role in climate change. This summer, the conservative news outlet Newsmax reported her as saying she did not think global warming was "man-made". In her ABC interview, though, Palin said: "Regardless, though, of the reason for climate change, whether it's entirely or wholly caused by man's activities or is part of the cyclical nature of our planet — the warming and the cooling trends — regardless of that, John McCain and I agree that we gotta do something about it."

How — or indeed if — McCain and Palin will resolve their large differences on science topics remains to be seen. In some areas the two have been moving closer; McCain has, for instance, come out in favour of offshore drilling for more oil and gas, a stance long favoured by Palin, as by most Alaskans. Stem cells, however, could turn out to be more of a stumbling block. As a senator, McCain twice voted to lift the federal restrictions on funding of human embryonic stem-cell research, but he has also spoken of adult stem-cell research and other new techniques, such as induced pluripotent stem cells, as an addition to funding the work on embryonic stem cells. In answer to a questionnaire submitted by the advocacy group Research!America, he wrote: "I also believe that clear lines should be drawn to reflect a refusal to sacrifice moral values and ethical principles for the sake of scientific progress." Last week, at a forum on health policy organized by the advocacy group Scientists and Engineers for America, McCain health-policy adviser Jay Khosla would not explicitly answer a question on the ban, reiterating McCain's prior voting record and interest in alternative approaches to human embryonic stem cells. "His real hope is that we can continue to develop new technologies," says Khosla.

Biomedical research advocates are starting to wonder what a McCain–Palin administration would mean for stem cells. "Although he has historically been somewhat independent from the White House on issues such as stem cells, I think with his vice-presidential pick, he would definitely not lift the ban," says Frankie Trull, president of the Foundation for Biomedical Research in Washington DC.

Top priorities

McCain advisers have been quick to point out that McCain, not Palin, is the presidential nominee, and as such his priorities take precedence. And it's possible that intense media focus on Palin since her nomination has exaggerated her possible influence on McCain's positions. She was not, after all, chosen for her background and expertise in policy. But her addition to the ticket is perhaps the most significant factor in recent months that could influence where science goes in the next administration.

As the campaign grinds its way towards its 4 November end, the science advisers for each campaign are still working to flesh out various platforms. Last month, for instance, McCain and Obama released detailed policies on space for the first time. Such policies are needed because of the electoral college votes up for grabs in the swing state of Florida, which stands to lose thousands of jobs when the space shuttle is retired in 2010, as current Bush policy holds. Clinton official Lori Garver helped to develop Obama's space platform, which calls for sending humans to Mars, whereas astronaut Walt Cunningham has represented the McCain camp at times on its space policies.

As they scramble to put together last-minute science advice, campaign representatives on both sides are also looking forward to life after November. One high priority is to draw up lists of possible names for appointments to top science positions in a new administration. The changeover from the Bush administration to a McCain or Obama one will be handled by a 'transition team', often incorporating some of the main campaign advisers and using them to help populate the new scene in Washington. "The most important thing about the transition is the president-elect's selection of the right team," says Neal Lane of Rice University in Houston, Texas, who served as a science adviser to Bill Clinton.

The next question will be how that new administration will work with the newly elected Congress; all 435 seats in the House of Representatives are up for grabs in November, along with one-third of the Senate seats. Still benefiting from the voter anger that swept them into power in 2006, the Democrats are widely expected to pick up seats in both the House of Representatives and the Senate. This could be crucial in the Senate, where Democrats govern by a narrow 51-49 majority, thanks to the fact that a pair of Independents has chosen to caucus with them.

> "The most important thing about the transition is the president-elect's selection of the right team."
> — Neal Lane

The delicate balance of Congressional power could be particularly important when it comes to climate-change legislation. In fact, Republican minorities in both the House and the Senate might be more inclined to work with a Republican president McCain on the issue. On the other hand, "if it's President Obama", says a Senate aide familiar with the discussions, "then many Republicans on the Hill will see it as their job to obstruct climate legislation simply because it is the priority of the president, who is a Democrat."

It's a salutory reminder that, whatever the next president's positions, their realization will depend on a lot of practical politics. And their implementation will depend on factors beyond even the president's power — from world food prices to medical breakthroughs to climate-linked catastrophes. For the next six weeks, though, it's all about who gets to face those uncertainties. ■

Alexandra Witze is *Nature*'s Chief of Correspondents for America. Additional reporting by Jeff Tollefson and Meredith Wadman.

Q&A

Questioning the candidates

Barack Obama accepted *Nature*'s invitation to answer 18 science-related questions in writing; John McCain's campaign declined. Obama's answers to many of the questions are printed here; answers to additional questions (on topics including biosecurity, the nuclear weapons laboratories and US participation in international projects) can be found at www.nature.com/uselection. Wherever possible, *Nature* has noted what McCain has said at other times on these topics.

What will be your highest priority in the science and technology arena?

Barack Obama: I am committed to strengthening US leadership in science, technology and innovation, and doing so will be a central priority for my administration. Our talent for innovation is still the envy of the world, but we face unprecedented challenges that demand new approaches. For example, the United States annually imports $53 billion more in advanced technology products than we export. China is now the world's number one high-technology exporter. This competitive situation may only worsen over time because the number of US students pursuing technical careers is declining. The United States ranks 17th among developed nations in the proportion of college students receiving degrees in science or engineering; we were in third place thirty years ago.

Federally supported basic research, aimed at understanding many features of nature — from the size of the Universe to subatomic particles, from the chemical reactions that support a living cell to interactions that sustain ecosystems — has been an essential feature of American life for over fifty years. While the outcomes of specific projects are never predictable, basic research has been a reliable source of new knowledge that has fuelled important developments in fields ranging from telecommunications to medicine, yielding remarkable rates of economic return and ensuring American leadership in industry, military power and higher education. I believe that continued investment in fundamental research is essential for ensuring healthier lives, better sources of energy, superior military capacity, and high-wage jobs for our nation's future.

Yet, today, we are clearly under-investing in research across the spectrum of scientific and engineering disciplines. Federal support for the physical sciences and engineering has been declining as a fraction of gross domestic product for decades, and, after a period of growth of the life sciences, the National Institutes of Health (NIH) budget has been steadily losing buying power for the past six years. As a result, our science agencies are often able to support no more than one in ten proposals that they receive, arresting the careers of our young scientists and blocking our ability to pursue many remarkable recent advances. Furthermore, in this environment, scientists are less likely to pursue the risky research that may lead to the most important breakthroughs.

D. DOVARGANES/AP

Finally, we are reducing support for science at a time when many other nations are increasing it, a situation that already threatens our leadership in many critical areas of science.

This situation is unacceptable. As president, I will increase funding for basic research in physical and life sciences, mathematics and engineering at a rate that would double basic research budgets over the next decade to support our scientists and restore US scientific leadership.

John McCain *has similarly promised to fight for increased funding at some of the country's leading science agencies, but without specific targets in terms of dollars or time frames. "Under a McCain administration, science and research will have a very high priority," says Jay Khosla, who advises the campaign on health policy issues. "He will do everything it takes to ensure we will continue to be leaders, especially in the field of innovation." One potentially major stumbling block to those looking for more research funding: McCain has said he would freeze domestic discretionary spending, which includes science money, for one year if elected in order to help trim overall spending levels.*

In general, McCain has stressed less government control and more business- and technology-oriented approaches to spurring innovation, such as a $300 million prize for advanced battery technology. On the campaign trail he touches on many of the same themes as Obama, such as the perceived need to educate more American scientists and engineers; McCain would, for instance, provide bonuses for high-performing teachers in subjects such as maths and science, and support education programmes at science agencies such as the National Science Foundation and the Department of Energy.

Biomedical innovation is expensive and very slow; it takes $1 billion and the better part of a decade to develop a new drug. What would your administration do to make it easier to turn research into cures?

Obama: Americans have good reasons to be proud of the extraordinary role that medical science has had in combating disease, here and throughout the world, over the past century. Work sponsored by the NIH, other government agencies and our pharmaceutical and biotechnology industries has produced many vaccines, drugs and hormones that have improved the quality of life, extended life expectancy and reduced the dire consequences of many serious illnesses and disabilities.

While it may never be easy to "turn research into cures", I understand that biomedical scientists are seeing enhanced opportunities to use their science to improve health. I will encourage the development of biological markers of disease that might simplify the evaluation of new therapies, the use of genetic information to select patients most likely to

L. LEFKOWITZ/GETTY

benefit from new treatments, and the multi-disciplinary efforts that are now possible at many research centres. In addition, I will support increased attention to research that focuses on prevention, early detection and improved management of disease.

Furthermore, I believe there is more that we can do to ensure new treatments are developed and made available to the public more efficiently. I believe that we must increase funding for the NIH to reverse the funding trends that have left our nation's scientists with fewer resources as research costs escalate. We must also do a better job of providing resources to the Food and Drug Administration (FDA), the Centers for Disease Control and Prevention (CDC), and other federal agencies that help ensure that when these medical advances are turned into exciting new treatments, we are able to ensure that they are swiftly and safely considered for widespread usage. Additionally, we must prioritize removing barriers both between federal agencies and across public, private and nonprofit organizations to ensure better and more efficient collaboration on new innovations.

McCain *has also said he would strongly support funding for the NIH. "McCain wants to make sure we are doing everything possible to give the young scientists of today the resources they need to go out and bring research to cures," says Khosla. The health-care platforms of both candidates tend to focus on how to make health insurance more affordable and accessible to Americans, and McCain has talked about technologies that could benefit public health — such as telemedicine — to bring the latest medical knowledge to many more patients.*

Many scientists are bitter about what they see as years of political interference in scientific decisions at federal agencies. What would you do to help restore impartial scientific advice in government?

Obama: Scientific and technological information is of growing importance to a range of issues. I believe such information must be expert and uncoloured by ideology. I will restore the basic principle that government decisions should be based on the best-available, scientifically valid evidence and not on the ideological predispositions of agency officials or political appointees. More broadly, I am committed to creating a transparent and connected democracy, using cutting-edge technologies to provide a new level of transparency, accountability and participation for America's citizens. Policies must be determined using a process that builds on the long tradition of open debate that has characterized progress in science, including review by individuals who might bring new information or contrasting views. I have already established an impressive team of science advisers, including several Nobel laureates, who are helping me to shape a robust science agenda for my administration.

In addition I will:

• Appoint individuals with strong science and technology backgrounds and reputations for integrity and objectivity to the growing number of senior management positions in which decisions must incorporate science and technology advice. These positions will be filled promptly with ethical, highly qualified individuals on a non-partisan basis;

• Establish the nation's first Chief Technology Officer (CTO) to ensure that our government and all its agencies have the right infrastructure, policies and services for the twenty-first century. The CTO will lead an interagency effort on best-in-class technologies, sharing of best practices and safeguarding of our networks;

• Strengthen the role of the President's Council of Advisors on Science and Technology (PCAST) by appointing experts who are charged to provide independent advice on critical issues in science and technology. The PCAST will once again be advisory to the president; and

• Restore the science integrity of government and restore transparency of decision-making by issuing an Executive Order establishing clear guidelines for the review and release of government publications, guaranteeing that results are released in a timely manner and not distorted by the ideological biases of political appointees. I will strengthen protection for 'whistle blowers' who report abuses of these processes.

McCain *has similarly pledged to fill key technical positions in his administration with qualified scientists and engineers, including having a science adviser working directly with the president. "McCain will seek to restore the credibility of scientific research" in the federal government, says campaign adviser Floyd DesChamps. McCain has argued that taxpayers' investment in scientific research should be repaid with the untarnished results of that work.*

What role does nuclear power have in your vision for the US energy supply, and how would you address the problem of nuclear waste?

Obama: Nuclear power represents an important part of our current energy mix. Nuclear also represents 70% of our non-carbon generated electricity. It is unlikely that we can meet our aggressive climate goals if we eliminate nuclear power as an option. However, before an expansion of nuclear power is considered, key issues must be addressed, including security of nuclear fuel and waste, waste storage and proliferation. The nuclear waste disposal efforts at Yucca Mountain [in Nevada] have been an expensive failure and should be abandoned. I will work with the industry and governors to develop a way to store nuclear waste safely while we pursue long-term solutions.

McCain *has proposed building 45 new nuclear power plants by 2030, with an eventual goal of a total of 100. McCain has not addressed where the nuclear waste from these and current reactors would go, and he has supported the Yucca Mountain storage project in the past.*

You support a cap-and-trade system for regulating greenhouse-gas emissions; what lessons from the European emissions-trading system would you implement?

Obama: I will implement a market-based cap-and-trade system to reduce carbon emissions by the amount scientists say is necessary: 80% below 1990 levels by 2050. While Europe has

E. SCHNAKENBERG/GETTY

SPL

had important successes with its system, it also has made mistakes that we should learn from. Unlike the European system, my plan would aim to cover virtually all greenhouse-gas emissions, would auction off all of the permits instead of giving them away, and would make sure there was stability in the market for permits and their price. My plan would use the proceeds of the auction for investments in a clean-energy future, habitat protection and rebates and other transition relief for families.

McCain *has described his own vision of a cap-and-trade system, but with a different target; the McCain plan calls for reductions of emissions by 60% below 1990 levels by 2050. McCain would initially give away emissions permits instead of auctioning them. McCain would also allow emissions allowances to be 'banked' or 'borrowed' for different time periods, as well as establish a national 'strategic carbon reserve' that could release permits during difficult economic times. He would also allow unlimited offsets, from both domestic and international sources, to ease into a newly set up cap-and-trade system.*

Does your stance on tapping domestic oil reserves stand at odds with your goals for reducing national emissions and combating climate change? How will you balance the two?

Obama: With 3% of the world's oil reserves, the United States cannot drill its way to energy security. But US oil and gas production plays an important role in our domestic economy and remains critical to prevent global energy prices from climbing even higher. There are several key opportunities to support increased US production of oil and gas that do not require opening up currently protected areas.

Increasing domestic oil and gas production in the ways I propose in no way lessens my commitment to combating climate change, one of the great challenges of our time. I am committed to implementing a market-based cap-and-trade system to reduce carbon emissions 80% below 1990 levels by 2050, and I will start reducing emissions immediately by establishing strong annual reduction targets with an intermediate goal of reducing emissions to 1990 levels by 2020.

McCain *currently favours a more aggressive offshore-drilling policy than Obama; both candidates, like the Democratic-led Congress, have changed their earlier stances opposing such drilling in the face of rising oil prices and public pressure to do something about it. However, McCain sees climate change as a national security issue, and maintains that it is a major priority for him. He emphasizes developing new emissions-reducing technologies with minimum costs in order to soften any blow to the national economy. McCain's intermediate goal for emission reductions is also 1990 levels by 2020.*

Do you believe that evolution by means of natural selection is a sufficient explanation for the variety and complexity of life on Earth? Should intelligent design, or some derivative thereof, be taught in science class in public schools?

Obama: I believe in evolution, and I support the strong consensus of the scientific community that evolution is scientifically validated. I do not believe it is helpful to our students to cloud discussions of science with non-scientific theories like intelligent design that are not subject to experimental scrutiny.

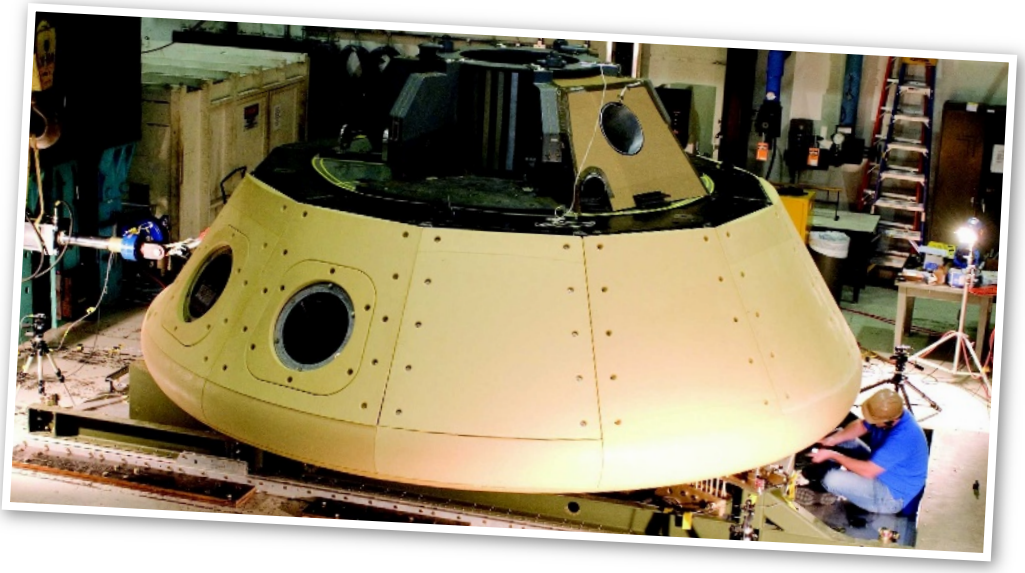
McCain *said last year, in a Republican primary debate: "I believe in evolution. But I also believe, when I hike the Grand Canyon and see it at sunset, that the hand of God is there also." In 2005, he told the Arizona Daily Star that he thought "all points of view" should be available to students studying the origins of humanity. But the next year a Colorado paper reported him saying that such viewpoints should not be taught in science class.*

Would you lift President Bush's ban on federal funding for research on human embryonic stem-cell lines derived after 9 August 2001? Under what conditions do you find it acceptable to create a human embryonic stem-cell line?

Obama: Stem-cell research holds the promise of improving our lives in at least three ways — by substituting normal cells for damaged cells to treat diabetes, Parkinson's disease, spinal-cord injury, heart failure and other disorders; by providing scientists with safe and convenient models of disease for drug development; and by helping to understand fundamental aspects of normal development and cell dysfunction.

For these reasons, I strongly support expanding research on stem cells. I believe that the restrictions that President Bush has placed on the funding of human embryonic stem-cell research have handcuffed our scientists and hindered our ability to compete with other nations. As president, I will lift the current administration's ban on federal funding of research on embryonic stem-cell lines created after 9 August 2001 through executive order, and I will ensure that all research on stem cells is conducted ethically and with rigorous oversight.

I recognize that some people object to government support of research that requires cells to be harvested from human embryos. However, hundreds of thousands of embryos stored in the United States in *in vitro* fertilization



NASA

ED KASHI/CORBIS

clinics will not be used for reproductive purposes, and will eventually be destroyed. I believe that it is ethical to use these extra embryos for research that could save lives when they are freely donated for that express purpose.

I am also aware that there have been suggestions that human stem cells of various types, derived from sources other than embryos, make the use of embryonic stem cells unnecessary. I don't agree. While adult stem cells, such as those harvested from blood or bone marrow, are already used for treatment of some diseases, they do not have the versatility of embryonic stem cells and cannot replace them. Recent discoveries indicate that adult skin cells can be reprogrammed to behave like stem cells; these are exciting findings that might in the future lead to an alternate source of highly versatile stem cells. However, embryonic stem cells remain the 'gold standard', and studies of all types of stem cells should continue in parallel for the foreseeable future.

Rather than restrict the funding of such research, I favour responsible oversight of it, in accordance with recent reports from the National Research Council (NRC). Recommendations from the NRC reports are already being followed by institutions that conduct human embryonic stem-cell research with funds from a variety of sources. An expanded, federally supported stem-cell research programme will encourage talented US scientists to engage in this important new field, will allow more effective oversight, and will signal to other countries our commitment to compete in this exciting area of medical research.

McCain'*s stance on embryonic stem-cell research has been the subject of much speculation among researchers. He has voted twice before to lift President Bush's funding restrictions on such work, but his running mate Sarah Palin opposes the work. His public position is perhaps best summarized in his response to questionnaires from advocacy groups such as Research!America last year and ScienceDebate2008 this year: "While I support federal funding for embryonic stem-cell research, I believe clear lines should be drawn that reflect a refusal to sacrifice moral values and ethical principles for the sake of scientific progress. Moreover, I believe that recent scientific breakthroughs raise the hope that one day this debate will be rendered academic. I also support funding for other research programmes, including amniotic fluid and adult stem-cell research which hold much scientific promise and do not involve the use of embryos. I oppose the intentional creation of human embryos for research purposes and I voted to ban the practice of 'fetal farming', making it a federal crime for researchers to use cells or fetal tissue from an embryo created for research purposes."*

Do you see astronauts on the Moon as a worthy goal for the country?

Obama: I believe that the United States needs a strong space programme to help maintain its superiority not only in space, but also here on earth in the realms of education, science, technology, the environment and national security. Technology developed for space missions has been applied to improve everything from computers and medical technology to baby formula and automobiles.

As president, I will establish a robust and balanced civilian space programme. In achieving this vision, I will reach out to include international partners and to engage the private sector to amplify NASA's reach. I believe that a revitalized NASA can help America maintain its innovation edge and contribute to American economic growth.

I will re-establish the National Aeronautics and Space Council, which will coordinate civilian, military, commercial and national security space activities and report to the president. This council will oversee a comprehensive and integrated strategy and policy dealing with all aspects of the government's space-related programmes, including those being managed by NASA, the Department of Defense, the National Reconnaissance Office, the Department of Commerce, the Department of Transportation and other federal agencies. It will solicit public participation, engage the international community and work toward a twenty-first-century vision of space that constantly pushes the envelope on new technologies as it pursues a balanced national portfolio that expands our reach into the heavens and improves life here on Earth.

Human spaceflight is important to America's political, economic, technological and scientific leadership. I will support renewed human exploration beyond low Earth orbit. I endorse the goal of sending human missions to the Moon by 2020, as a precursor in an orderly progression to missions to more distant destinations, including Mars.

McCain *has released an extensive space platform, including calling space exploration a "top priority" for the country and manned space flight "a reflection of national power and pride". Unlike Obama, McCain has explicitly committed to funding the Constellation programme to replace the space-shuttle fleet (although without details on how he would accomplish that). He also says he would maintain the nation's space infrastructure, including the related workforce, and focus on maximizing the research possibilities of the International Space Station. He would maintain investments in aeronautics research as well as the infrastructure for Earth-monitoring satellites.*

Would it make sense for more overseas students who receive PhDs at American universities to stay in the country and contribute to its research base and its wealth? What immigration reforms would you support?

Obama: I believe that we must enact comprehensive immigration reform to restore our economic strength, relieve local governments of unfair burdens stemming from an inefficient federal immigration system, ensure that our country and borders remain secure and allow a path to citizenship for the 12 million undocumented immigrants who are willing to pay a fine, pay taxes, and learn English. A critical part of comprehensive immigration reform is turning back misguided policies that since 9/11 have turned away the world's best and brightest from America. As president, I will improve our legal permanent resident visa programmes and temporary programmes to attract some of the world's most talented people to America.

McCain, *as a senator from Arizona, has long been involved in immigration issues, mainly through strengthening federal security at land border crossings. He supports immigration reforms to allow more highly skilled workers to stay and work in the United States after graduation.* ■

Reporting by Alexandra Witze.

See Editorial, page 431.

Agencies of change

A new president could bring radical shifts to America's major research entities. *Nature* profiles some of the agencies in need of a makeover.

Going for the stars — or going broke?

When **NASA's** rockets shot astronauts to the Moon, the Apollo programme had the backing of a nation — and its chequebook. At its peak in 1966, NASA was spending just under $6 billion a year. That was 4% of the federal budget at the time, and nearly $38 billion in today's money.

Some things are similar today. Four years ago, President George W. Bush revived the notion of a Moon shot with his Vision for Space Exploration. And although it's not quite the space race of the 1960s, international competition is heating up again, thanks to heightened tensions with Russia over the Georgia conflict plus nascent competition with China, which plans to conduct its first spacewalk shortly.

But the main difference is money. Today, NASA's budget is $17.3 billion, less than 0.6% of federal spending. NASA is now in a squeeze as it tries to develop its next Moon rockets, a programme called Constellation that is scheduled to lift astronauts into orbit by 2015, and to the Moon by 2020. When either John McCain or Barack Obama takes office next year, he will have simple alternatives, says Len Fisk of the University of Michigan in Ann Arbor. "You've got two choices: more money or less programme," says Fisk, a former NASA science chief and former chair of the National Academies' Space Studies Board.

NASA administrator Mike Griffin has tried to find more money for Constellation by holding firm to a 2010 retirement for the space shuttle fleet, which costs more than $3 billion annually. From 2010 until 2015, when the Constellation rocket would theoretically be ready, NASA would have to buy rides aboard Russia's Soyuz spacecraft to ferry astronauts to the International Space Station. The gap — already deemed "unseemly" by Griffin — has become even more prominent in an election year in which Florida, a crucial swing state, has thousands of shuttle-related jobs at stake.

Obama has said he would support additional shuttle flights, and a $2-billion boost to NASA's budget to that end. McCain, along with two other senators, recently sent a letter to the White House imploring the Bush administration to leave open the option of extending shuttle flights beyond 2010.

But if campaign promises don't materialize in congressional spending bills, the agency will have to find other ways to squeeze more out of a budget that has remained roughly flat in real terms for the past two decades. One way would be to improve accountability in its contracts, suggests planetary scientist Alan Stern, who left NASA in March after a one-year stint as science chief.

Within the science mission directorate alone, which receives roughly a quarter of NASA's budget, 10 missions had collectively been delayed by 85 months and run up $430 million in unanticipated costs within the past two years. "We have weather satellites that cost $3.5 billion. This is absurd. We have Mars missions that have tripled in cost and no one blinks an eye," says Stern. "You reward those who do not show discipline at the expense of those who did."

Eric Hand

A new rocket, called Ares, is under development in the Constellation programme.

NASA/MSFC

Curing the country

The 2008 presidential election comes at a critical juncture for the **National Institutes of Health** (NIH). From 1998 to 2003, the biomedical agency's budget doubled to $27 billion; in the five years since, its purchasing power has eroded as congressional increases have failed to keep pace with inflation. The transition from rapid growth to freeze has shaken the nation's biomedical enterprise, causing senior people to face odds as low as one in ten of getting funding.

Elias Zerhouni, the NIH director, this month singled out funding difficulties as hampering the agency's long-term planning. Like all federal agencies, the NIH gets its money through annual congressional appropriations, and the numbers cannot be depended on from year to year. "The way the process works does not allow us to do good medium- and long-term capital investments," Zerhouni told a congressional subcommittee.

A new president won't be able to change the way Congress funds agencies, but he might have something to say about increasing the total bottom line for the NIH. Both leading candidates have come out in favour of biomedical research spending, with subtle but possibly important differences between the two.

In a 2007 questionnaire by advocacy group Research!America, Barack Obama wrote: "I strongly support increasing funding for the NIH." John McCain responded to the same question with the ever so slightly different words: "I strongly support funding for the NIH."

J. WATSON/AFP/GETTY IMAGES

Elias Zerhouni wants predictable funding levels.

To NIH supporters, the writing on the wall seems, if not crystal clear, at least legible. "It's clear that Senator Obama will act to increase funding for the NIH. We do not see that in the McCain response," says Mary Woolley, president of Research!America, which is based in Alexandria, Virginia. Woolley nonetheless believes that the force of public opinion is so strongly in favour of medical research that McCain will deliver on an NIH increase.

Other NIH advocates are similarly positive. "We are very optimistic that a new administration, regardless of which party is elected, will support renewed investment in the NIH and other science agencies," says Richard Marchase, president of the Federation of American Societies for Experimental Biology and vice-president of research at the University of Alabama, Birmingham.

That may be wishful thinking in an era of blooming US budget deficits and growing pressure on Congress to curb spending. Even with presidential backing, cooperative lawmakers — including NIH advocates and purse-string holders Senator Tom Harkin (Democrat, Iowa) and Representative David Obey (Democrat, Wisconsin) — may not be able to deliver NIH increases in the face of multiple, competing priorities.

And McCain has floated the idea of a one-year spending freeze across the board for all agencies to help cope with the ballooning federal deficit.

Adding to the morose budgetary atmosphere is a ban on federal funding for research done on human embryonic stem cells lines generated after 9 August 2001. US researchers have watched their international colleagues push ahead in what is arguably one of the most promising areas of biomedical research.

In the most recent fiscal year, only $42 million of the $650 million spent by the NIH on stem-cell research went to work on human embryonic lines. And although many expect the ban, put in place by President Bush, to be lifted no matter who is elected — McCain and Obama have both voted in favour of lifting the ban in the past — the choice of conservative Sarah Palin as McCain's running mate has some researchers wondering what that might signal (see page 442).

Meredith Wadman

Cleaning up the mess

Environmentalists and scientists alike are hoping that the next president will set the **Environmental Protection Agency** (EPA) on a new course.

Under President Bush, the agency has struggled against accusations of political interference in scientific decisions dating back to 2002 when the White House reportedly meddled with the EPA's assessments on global warming. By July this year, things at the agency had deteriorated to the point that leading Senate Democrats were calling on its chief, Stephen Johnson, to resign.

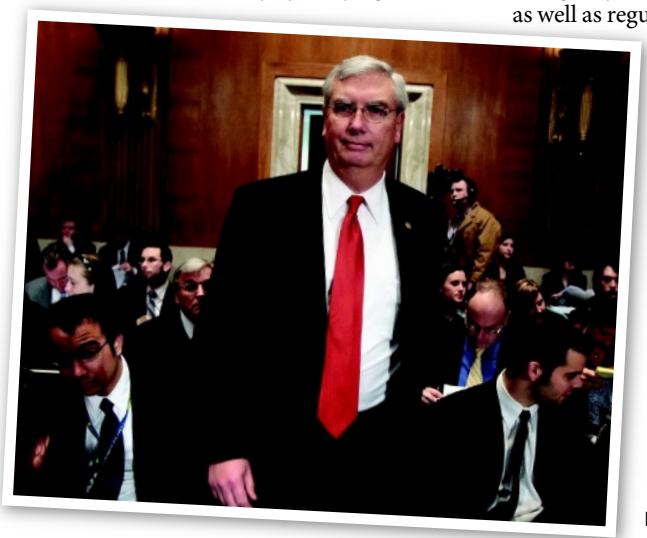
The main fight is over global warming. In 2007, the US Supreme Court ruled that the EPA had authority to regulate carbon dioxide pollution from vehicles. Johnson, though, denied California its application to separately regulate CO_2, saying the problem should be addressed at the federal, not local, level. In that decision he overruled his staff, reportedly after receiving pressure from the White House to do so. He then followed up in July by delaying action on the agency's own regulatory proposal, insisting the issue of regulating greenhouse gases was better left to Congress.

Many experts agree that Congress is the only real forum for a deal between environmental and industrial interests, but environmentalists and some politicians say that the EPA should take action while that debate is under way. Indeed, some see EPA regulations as a tool to force action in Congress, because industry interests would rather have an open debate about legislation. Moving forward, the question is how either presidential candidate would work with Congress, the legislative body, and the EPA, the regulatory body, to set up a greenhouse-gas programme.

Obama's advisers are quick to suggest their candidate would use both legislative and regulatory approaches to help control emissions. "If the circumstances are not right to get that legislation passed quickly, then I think the regulatory authority would come to the fore," says Robert Sussman, a Democratic adviser who served as deputy EPA administrator under Bill Clinton. McCain's campaign has been quiet on the issue, but many experts say the Republican would be less likely to take the regulatory route.

Regardless, the next EPA administrator may find his or her hand forced by lawsuits. Just last month, New York and 11 other states sued the EPA for failing to regulate CO_2 emissions from oil refineries.

But the alleged problems run much deeper than any given policy position. Reports of political interference in agency decisions are widespread, and many say they have contributed to morale problems throughout the agency. Union leaders representing EPA scientists blasted Johnson earlier this year, criticizing the agency's proposal to close several libraries as well as regulatory decisions regarding pesticides, drinking-water standards and mercury emissions from power plants.



C. SOMODEVILLA/GETTY IMAGES

EPA chief Stephen Johnson, under fire.

Jeff Ruch, executive director of Public Employees for Environmental Responsibility in Washington DC, says it's not yet clear what kind of management style either Obama or McCain would bring to the agency. "The question is the extent to which a president will pledge to let the scientists talk," he says. "I have not yet heard that pledge by either candidate." ■

Jeff Tollefson

For more on the US election, see www.nature.com/uselection.

Not the best advice

Concerns about the next president's science adviser miss the real issues, says **David Goldston**.

Two prestigious groups recently issued recommendations on how to strengthen the science advisory arm of the White House — the Office of Science and Technology Policy (OSTP). Most of the suggestions are pretty basic; it's hard to imagine anyone disagreeing with them, at least in the abstract. But they raise a more fundamental question: should the science community be so obsessed with the OSTP?

The recommendations come from the Woodrow Wilson International Center for Scholars[1] and the Center for the Study of the Presidency[2]. They fundamentally offer the same advice — appoint a prestigious scientist to head the OSTP early in the next administration, fund the office well, and ensure that the president hears from many scientists. These are all old chestnuts even if they haven't always been followed.

But a more pointed recommendation in the two reports highlights how the focus on the OSTP may say more about the mindset of scientists than about the nature of government. Both groups begin by arguing that the OSTP director should also hold the title of Assistant to the President for Science and Technology. This was always the case until George W. Bush took office. In fact, the presidential assistant position existed decades before the statutory creation of the OSTP in 1976. The loss of the title has stuck in the craw of science-policy panjandrums for eight years because it seemed to signify that scientists had been cast out of the president's inner circle.

It would indeed be a welcome gesture for the next president to restore the title, but would it make any practical difference? Can anyone cite any decision that has been different because the current head of the OSTP, John Marburger, was not called 'Assistant to the President'? The prominence given to the recommendation about a title speaks volumes about the scientific community's hypersensitivity to perceived slights and its excessive insecurity about its stature, but it says almost nothing about governance.

Although the Bush administration has never explained why it withdrew the title, it seems to be part of this White House's general effort to reveal as little as possible about the inner workings of government. The position of director of the OSTP was created by Congress, and the appointment must be confirmed by the Senate.

PARTY OF ONE

It is more difficult legally for a president to make claims of confidentiality for advice given by an official with a congressional pedigree. Generally, presidential assistants are not beholden to Congress in this way; they are simply the president's personal advisers. The last thing the Bush administration would have wanted was to give the title of presidential assistant to an official over whom Congress could exercise some authority.

This isn't just a theoretical concern. In 2005, the Bush administration went so far as to try to prevent Floyd Kvamme, a venture capitalist who chairs the President's Council of Advisors on Science and Technology (PCAST), from testifying before Congress, arguing that he was a presidential adviser whose views should be confidential. The White House eventually backed off, partly based on the argument that the advisory group was congressionally chartered. As a congressional staffer, I helped lead the battle against the White House in this dispute, which never became public.

The science community is blind to all this because of its insular focus. It tends to assume that decisions related to science policy primarily reflect attitudes towards scientists and science when in fact they are often driven by broader concerns. As a result, the two reports implicitly asked the wrong question about a president's politics. The best indicator about the future OSTP director's title may be a candidate's views on government secrecy, not science.

The focus on the OSTP itself may be similarly misdirected. Certainly the OSTP can be important in shaping policy, and its more mundane role of coordinating inter-agency research efforts is essential. But is it the best place to start an inquiry into the future of science policy? The reports seem to assume that having a well-known science adviser with good access to the president will mean scientists will be happy for the next four years. But that just isn't the case.

Both documents look back fondly on Allan Bromley, the late Yale University physicist who ran the OSTP during the administration of the elder George Bush. Bromley had a long-standing personal relationship with the president, and even got Bush to invite PCAST to Camp David, the presidential retreat. But that didn't make that first Bush administration some kind of golden age of science.

A major science issue at the time was the extent to which the federal government should invest in research to strengthen US competitiveness, then taking a beating from Japan. The administration of the elder Bush remained internally divided on the issue, with the opposition to such investments led by the director of the Office of Management and Budget — a far more powerful post than OSTP director even though it doesn't carry the title of presidential adviser. As a result, the federal government had made only tepid movement towards a competitiveness policy by the time Bush left office in 1993.

Neither Bromley's stature as a scientist nor his closeness to the president enabled him to control the administration's position on research spending, which was not determined by the president's attitude towards science. The issue came down to a philosophical debate about the role of the government in the economy. The science budget was affected, as always, by factors far broader than science and therefore by officials with broader portfolios.

Other interest groups tend not to fall into this trap of worrying about prestige and narrowly defined issues. Environmentalists, for example, don't spend much time writing recommendations about the White House Council on Environmental Quality, an office with a parallel role to that of the OSTP. And they don't just focus on specific policy goals; they worry if a candidate's overall governing philosophy indicates that little will come of his professed warm feelings towards the environment.

The science community's focus on status, access and love of science can cause it to overlook the broader issues that are more likely to determine how a president treats scientists. ■

David Goldston is a project director at the Bipartisan Policy Center in Washington DC. Reach him at partyofone@gmail.com.

1. http://wilsoncenter.org/news/docs/OSTP%20Paper1.pdf (2008).
2. http://election2008.aaas.org/docs/CSP_PresidentialLeadership.pdf (2008).

For more on the US election, see www.nature.com/uselection

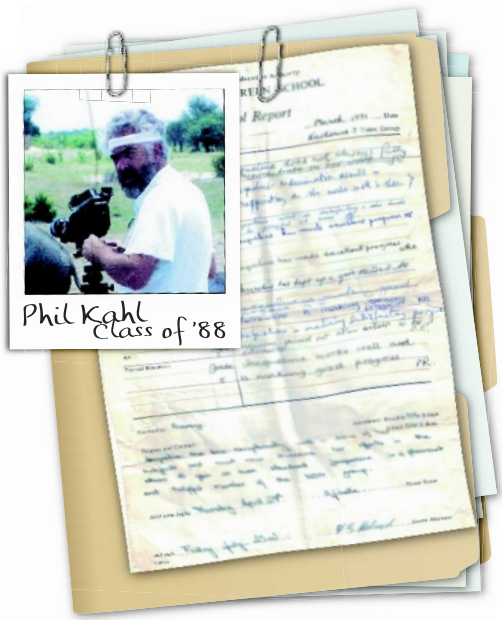
Best in class

What is it like to be labelled a genius? **Kendall Powell** follows the paths of four MacArthur Fellows — and finds they lead to rutting elephant bulls, climate-change champions, hybrid sunflowers and robotic hands.

In what is perhaps the most secretive talent contest ever, the MacArthur Foundation recognizes and rewards people who demonstrate exceptional creativity. Just this week, the foundation awarded 25 disbelieving recipients in the arts, humanities and sciences with its illustrious MacArthur Fellowships or, as the media has labelled them, 'genius grants'.

The fellowships don't have quite the intellectual and financial cachet of a 10-million Swedish krona (US$1.5-million) Nobel prize. But they do bring with them the label of 'genius', and half a million dollars to spend however the recipient decides. Now in its 28th year, the programme has awarded more than $360 million to 781 US citizens and residents aged from 18 to 82.

The MacArthur Foundation, based in Chicago, Illinois, does not give the awards for specific accomplishments. Through a highly secretive nomination process — files are allowed to 'ripen' over several years — the programme identifies extremely talented and creative people who are likely to make breakthroughs and make a lasting contribution to society. "We get the best information we can and consult with the wisest people we know," says Mark Fitzsimmons, associate director of the MacArthur Fellows programme and the programme's in-house scientist. "We think it is a good investment to build up the portfolio of young people who can change their field significantly."



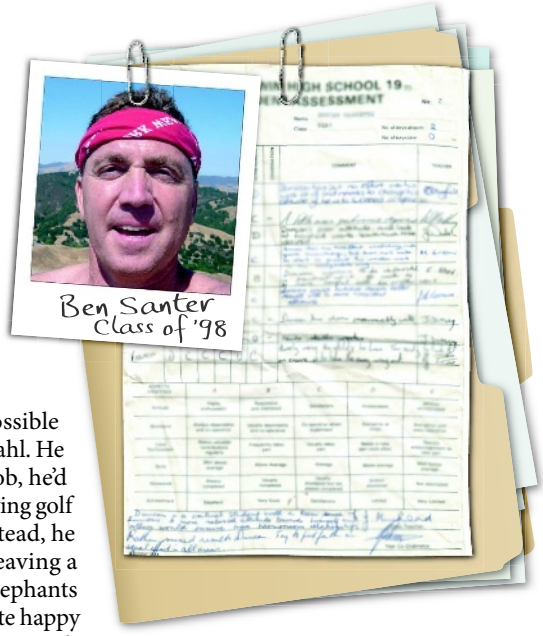
ARTWORK BY W. FERNANDES

It is also an investment with no guaranteed return: the foundation makes no attempt to measure the productivity of fellows or even what career path they follow. So do the geniuses go on to scale new intellectual heights, or do they struggle with the weight of expectation? *Nature* catches up with four individuals who got 'the call' 1, 5, 10 and 20 years ago and who have used their awards to push forward the frontiers of science.

Class of 1988: Philip Kahl

When Philip Kahl won a MacArthur award in 1988, he was already known as a zoologist who followed his own path. He had spent 35 years studying bird behaviour and he had done it all as a freelancer.

Even though colleagues pressured him to settle into a 'real job', Kahl had pieced together grants from the National Science Foundation, National Geographic and the National Audubon Society to travel to exotic



field sites. He also funded his work by selling his wildlife photography, including a 1969 *National Geographic* cover of an elephant in the Kenyan savannah. "I was having a helluva lot more fun than if I had been teaching at a university," says Kahl, now 73, from his home in Sedona, Arizona.

With the MacArthur money, Kahl's income "quadrupled overnight". Through investment and simple living, his fellowship (US$320,000 in 1988) has grown into an endowment that has largely sustained him and his work ever since. It also gave him the opportunity to switch his studies to the elephants he had been photographing.

"The birds were great and very photogenic — but you can't get attached to birds, they are too reptilian," he says. "Elephants you can really grow to love and still stay objective." Kahl says that no one else would have given him a grant to study a topic that, on paper at least, he knew little about. He used some of the money to spend six seasons in the 1990s studying African elephants' visual communication displays in Zimbabwe's Hwange National Park.

Since then, Kahl has been working at home analysing 225 hours of digital videotape frame by frame, and writing a book that catalogues the 100 or so visual displays elephants exhibit in the wild. He does not toil in obscurity. Elephant researchers around the world recognize his extreme dedication and deep knowledge of elephant displays. He collaborates by e-mail and phone and because he no longer travels, colleagues go to visit him and his library of 3,800 scholarly papers on elephants in about a dozen filing cabinets in his home office.

B. EDWARDS/THE ADVERTISER

"I can only take a week at a time with Phil," says Bob Dale, who studies animal behaviour at Butler University and the Indianapolis Zoo in Indiana. "Because it's breakfast, then work on elephants, lunch, work on elephants, dinner, more elephants."

This focus has helped Kahl make observations about the animals that had been missed before. He discovered that 'rutting' African musth bulls can be identified by a wrinkle about two-thirds of the way down their trunks. He also documented how only older musth bulls mate during the peak rainy season; younger bulls must wait until the off season[1]. His observations of these and other displays are valuable for conservationists and zookeepers who want to better understand elephants in the wild and in captivity. (Kahl makes no secret that he identifies with these testosterone-driven males — his e-mail moniker is 'musthbull').

"Phil has no time for the pretensions of the academic society," says Dale. "But he admires an intellect wherever he finds it."

"Many people don't think it's possible to be a freelance researcher," says Kahl. He says that if he had ever got a real job, he'd most likely be retired now and "playing golf or some other useless pursuit". Instead, he works everyday with the goal of leaving a legacy of some knowledge about elephants that will benefit elephants. "I'm quite happy I never got a job. A lot of people who get the fellowships don't do diddly, they spend the money on a house or something. I would like to think I did a lot more than diddly. I just hope I don't kick off before the book is done."

Class of 1998: Benjamin Santer

Benjamin Santer listens to National Public Radio (NPR) on his morning drive to work. And every day, when the announcer says, "Support for NPR is provided by the John D. and Catherine T. MacArthur Foundation," he lets out a little "Woohoo!" Ten years ago, Santer was at an utterly gloomy point, both professionally and personally. The call from the MacArthur Foundation, "was like a miracle, it changed my life", he says.

"The call from the MacArthur Foundation was like a miracle, it changed my life." — Benjamin Santer

Santer was in the middle of a bitter legal fight for custody of his four-year-old son and the fees were forcing him into debt. To remain near his son, he was seriously considering resigning from the Lawrence Livermore National Laboratory in California, where he did statistical analysis of climate models.

He was also on the stand at work. He had spent the past two years defending a single sentence that he had penned in the 1995 Intergovernmental Panel on Climate Change (IPCC) Second Assessment Report. Those 12 historic words — "The balance of evidence suggests a discernible human influence on global climate" — became one of the report's major conclusions and unleashed a firestorm around Santer, including calls for his dismissal and editorials in national papers denouncing his science. Energy, transportation and industrial producers of greenhouse gases were keen to avoid blame and emissions curbs.

Although 35 contributors agreed to the wording, Santer was singled out because he was the lead author of the relevant chapter and much of his own research had helped to build the case that some global warming must be attributed to human activities. One of his key studies showed that an anthropogenic 'fingerprint' on climate change predicted by computer models could be detected in weather-balloon data[2], and was cited as "the most convincing demonstration yet" of a human contribution to changing global air temperatures[3].

"Every day I was getting up and putting on this battle armour," recalls Santer. "The recognition that the MacArthur brought showed me that I wasn't alone, that there were others who thought this battle to preserve the integrity of science against these powerful interests was worth fighting." The award money allowed him to pay his legal fees, refocus his energy on science and helped secure a major grant from the Department of Energy to pursue his academic studies. (In 1999, he won custody of his son.) Santer says that the MacArthur prestige also generated its own pressure that has followed him for the rest of his career. "People expect you to do extraordinary things and for pearls of wisdom to come out every time you open your mouth," he says.

Santer went on to show that the human fingerprint was evident in many other parts of the climate system, such as atmospheric water vapour, the height of the tropopause (the upper limit of the weather-containing layer of atmosphere) and ocean surface temperatures in regions where hurricanes form.

He and his colleagues also tackled another of the field's controversies — why a satellite data

set developed at the University of Alabama in Huntsville showed cooling temperatures in the atmosphere in the tropics, whereas computer models invariably predicted warming in response to greenhouse gases. The apparent discrepancy was a hot issue for more than a decade because it allowed climate-change sceptics to question the validity of the various climate models.

Encouraging the next generation's creativity, Rieseberg paints with his son Mendel, named after the geneticist.

M. TSENG

In 2005, scientists at Remote Sensing Systems in Santa Rosa, California, found that the cooling results were due to an error in correcting the Alabama data for the effects of satellite orbital drift and created a new data set that was adjusted correctly. Led by Santer, a large team including many high-profile climate scientists, then showed that the new data set was in good agreement with computer models[4].

Only the most extreme sceptics now doubt that humans have contributed to the global climate. "I have felt the responsibility to continue to do science that helps us understand the magnitude of the problem," says Santer, who has continued to contribute to subsequent IPCC reports.

Karl Taylor, a climate scientist and longtime collaborator with Santer at Lawrence Livermore, says that Santer's 'genius' lies in the ability to coordinate group efforts to address important questions. "He's a driver of work and he has the confidence of the scientific community. When he asks, people are willing to jump in and help," he says.

Like most other people awarded with a MacArthur fellowship, Santer is uncomfortable with the label of 'genius'. "I don't think of myself as a genius," he says. "If I ever develop any kind of ego or maniacal characteristics, my friends and family should feel free to take the nearest blunt object and whack me on the head."

Class of 2003: Loren Rieseberg

Loren Rieseberg isn't a science superstar. He has no fancy educational pedigree. He doesn't exhibit eccentricities or ego. "On the outside, he is exceedingly calm, very soft-spoken, and kind," says Gerry Gastony, a plant systematist working with Rieseberg at Indiana University in Bloomington. "All that belies the intensity within. He has a laser-like focus."

As an undergraduate, Rieseberg attended Southern Adventist University in Collegedale, Tennessee, where the teaching of evolution was banned. To get around the restrictions, the biology department taught a class called 'speciation' because as Rieseberg puts it, even Bible college types "don't mind a little micro-evolution".

Captivated by the subject, Rieseberg pursued a PhD in plant speciation at Washington State University in Pullman. "I was canny enough to realize even at that early stage that grant funding would be easier if I was working on a wild relative of a crop plant — and sunflowers are the one of very few crop plants to be domesticated in North America," says Rieseberg. "Plus, I like the way they look."

Rieseberg turned down posts at Harvard University in Cambridge and at the University of Michigan in Ann Arbor for a job at Rancho Santa Ana Botanic Garden in Claremont, California, where, he says "there were no committees to sit on". It was there in the early 1990s that he started the work that most likely opened his MacArthur file, showing that new species of sunflowers can arise from hybridization.

N. KANE

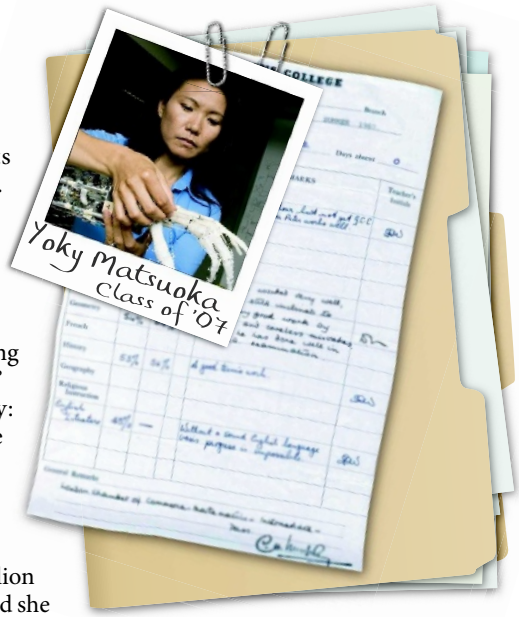
Hybrids are often considered an evolutionary dead-end, because they tend to have low or no fertility: think mules, the generally sterile offspring of male donkeys and female horses. Rieseberg investigated how new, viable species could arise by crossing together two US sunflower species to make experimental hybrids and then using nearly 200 genetic markers to roughly map their genomes. The work showed that the greenhouse-made hybrids had a similar genetic make-up to an ancient, natural sunflower hybrid species, suggesting that selection rather than chance determines which genetic combinations survive as hybrid species[5].

The genetic mapping was a technical tour de force, taking almost 8 years and 100,000 manual polymerase chain reactions. And the study changed the way that botanists and zoologists think about hybridization, says Jeannette Whitton, who was a postdoc with Rieseberg and is now an evolutionary biologist at University of British Columbia in Vancouver. "When you form a hybrid, yes, the first generation's genome is 50–50 between the parental species. But after that, all bets are off — selection pulls advantageous genes through to the following generations and disadvantageous genes get left behind," says Whitton

Rieseberg says that winning the MacArthur fellowship "has made everything easier". He says the biggest impact was in greasing the wheels for the next grants — a comment repeated by other winners. It helped him expand his group, which at around 20 people is arguably the largest plant evolutionary biology lab in North America; in the five years since his MacArthur fellowship, Reiseberg has published 50 original research papers. The influx of cash allowed Rieseberg to maintain his laboratory at Indiana and served as a 'piggybank' to set up another at the University of British Columbia, which is close to where he grew up.

Since the MacArthur award, Rieseberg and his group have solved a long-running debate by showing that sunflowers were first domesticated on the east coast of the United States rather than in Mexico, as some had thought. They also showed that introducing a genetic modification for pest resistance is highly advantageous and likely to spread from crop sunflowers to wild relatives[6].

Rieseberg thinks that the aura surrounding the MacArthur award, rather than the money itself, has made the biggest difference to him in his career. He suspects that the MacArthur Foundation earned its reputation by selecting people from all ages and disciplines and by rewarding people early in their careers. "There are a handful of MacArthur Fellows who are truly head and shoulders above everyone in their field," he says. "But a lot more of us are lucky."



M. LEVIN, UNIV. WASHINGTON PHOTOGRAPHY

Class of 2007: Yoky Matsuoka

All the MacArthur award recipients remember the moment they got the call. When neuroscientist Yoky Matsuoka picked up the phone last September, she assumed the caller was a prankster. "I'm going to tell you something very shocking," said the voice. "Are you sitting down? If you are holding anything precious, like a baby, please put it down."

Matsuoka was indeed holding a baby: she'd given birth eight days earlier. The caller, programme director Daniel Socolow, promised to try again in half an hour. When he did, he told her she had been named a MacArthur Foundation Fellow, that she would receive half a million dollars with no further obligation and did she have any questions? "I'll never forget in my life how it happened," says Matsuoka. "And one year later, it still doesn't feel real."

For Matsuoka, who works in robotics at the University of Washington in Seattle, the sum of money could be dwarfed by larger research budgets. "If I just let it sit, it could easily get absorbed into daily things," she says. Instead, Matsuoka has spent the past year pondering how to do something special with her fellowship. "In academic life, it's more about buying time than anything else," she says. And time is at a premium for this 37-year-old whose stated lab and life motto is 'work hard, play hard'.

While in graduate school at the Massachusetts Institute of Technology in Cambridge, Matsuoka became frustrated by the lack of neuroscience in the field of robotics and went on to combine the two in the emerging field called neurobotics. Her laboratory's main goal is to produce the Anatomically Correct Testbed Hand — a prosthetic that resembles and functions like a human hand, including synthetic parts for the bones, tendons, muscles and skin. It would be controlled by the same brain signals that control a natural hand. The aim is to provide rehabilitation and assistance to people with disabilities, and to understand how all the moving parts and neural signals generate such dexterity.

So far, Matsuoka has developed a robotic hand with three of the fingers moved by several motors, representing the muscles that give the human finger its many ranges of motion[7]. Current prosthetics, by contrast, typically offer only one range of motion, opening and closing the fingers. Her group has also tested whether 'molecular wires', polymers designed to conduct electricity across cell membranes, can be inserted into neuron membranes to connect a silicon prosthesis directly to existing nerve cells[8]. The team has succeeded in artificial cell membranes but has yet to translate their findings to real neurons.

"Yoky can do pretty much whatever she wants," says Matt O'Donnell, the dean of engineering who helped recruit Matsuoka to the University of Washington in 2006. He says the synthesis of living tissue engineering with non-biological components in the future is likely to overtake a purely robotic project such as the testbed hand, but that if anyone can do this type of synthesis, it's Matsuoka.

Fitzsimmons says that, like many recipients, Matsuoka was chosen because there was evidence, including testimony from her peers, of her creativity and diligence. She was selected "with the confidence that she can and will make important strides in whatever direction she pursues", he says.

M. LEVIN, UNIV. WASHINGTON PHOTOGRAPHY

Since receiving the call, Matsuoka has decided to use the MacArthur money to make a more immediate difference in the lives of people with disabilities. Her inspiration was a man she had met about five years ago who complained that the umbrella attachment for his wheelchair only allowed for a vertical position, whereas rain comes down in all directions. She immediately recognized this as a simple engineering problem — one easily solved by undergraduate or graduate students. She figured there must be many such ways that engineering expertise could improve the quality of life for people with disabilities, but how would she organize such an effort on top of her already demanding duties?

Enter Rayna Liekweg, a former manager for IBM with a background in mathematics, who had spent the past decade raising her children in the Seattle suburb of Kirkland. Liekweg read about Matsuoka being tapped as a MacArthur Fellow in the newspaper and, fascinated by her neurobotic work, cold-called her to see if there was a way she could help as a volunteer. "There is a whole community of women out there not unlike me," says Liekweg. "We are college-educated, have good skills and are looking for something meaningful and intellectually challenging to contribute to."

Matsuoka is in the process of forming a non-profit organization to draw on this community of women, making them into project managers who would connect a problem encountered by people with disabilities with engineering students who could build solutions. As a mother of three children under three-years-old, Matsuoka says she is passionate about helping and encouraging women who are balancing work and family or who gave up their careers.

She also wants to change the image of science and maths for young girls so that being captain of the National Science Bowl team becomes as hip as being captain of the girls' soccer team. "Now that I'm in my thirties, it's time for me to be disseminating what I did wrong and what I did right to those in their teens and twenties."

Matsuoka admits that she was worried at first about getting stuck with the label 'genius' and how it might fuel jealousy. She had heard the sniping comments colleagues made behind the back of a researcher who had won another prestigious award. "But after the first three months I let that go. People are going to talk. I have to do the productive things I can do because of it. Then they can say, 'Maybe she's not a genius, but at least she's done something with it." ■

Kendall Powell is a freelance science writer based in Broomfield, Colorado.

1. Kahl, M. P. & Armstrong, B. D. *Mammalia* **66,** 159–171 (2002).
2. Santer, B. D. *et al. Nature* **382,** 39–46 (1996).
3. Nicholls, N. *Nature* **382,** 27–28 (1996).
4. Santer, B. D. *et al. Science* **309,** 1551–1556 (2005).
5. Rieseberg, L .H., Sinervo, B., Linder, C. R., Ungever, M. C. & Arias, D. M. *Science* **272,** 741–745 (1996).
6. Harter, A. V. *et al. Nature* **430,** 201–205 (2004).
7. Vande Weghe, M., Rogers, M., Weissert, M. & Matsuoka, Y. *Proc. IEEE Intl Conf. Robot. Automat.* **4,** 3375–3379 (2004).
8. Widge, A. S., Jefferies-El, M., Cui, X., Lagenaur, C. F. & Matsuoka, Y. *Biosens. Bioelectron.* **22,** 1723–1732 (2007).

CORRESPONDENCE

Shaping science education in just 100 words

SIR — A science workshop held in Venice earlier this year under the banner of '100 parole per la scienza' challenged a group of one hundred 16–18-year-olds to choose 100 words that, in their collective opinion, represent crucial factors and concepts influencing trends in science today. The students were from schools all over Italy and the workshop was organized by the San Paolo Fondazione per la Scuola and Fondazione Venezia (www.100parole.it).

Their final list was assembled after an imaginative range of seminars from notable scientists and thinkers, and after extensive discussion and individual word searches of scientific works on the web and in books and journals. Here is the result, in alphabetical order: Acid/base, aggregation status, analysis, antimatter, apparatus, atmosphere, atom, bacteria, Big Bang, biodiversity, bioethics, biosphere, black hole, carbon, cell, chaos, climate, cloning, DNA, ecosystem, electricity, electron, element, energy, entropy, environment, enzyme, equilibrium, error, ethology, evolution, experiment, force, fossil, galaxy, gene, genetically modified organism, gravity, greenhouse effect, H_2O, heat, hydrocarbon, infinity, intelligence, Internet, life, light, link, magnetism, mass, matter, measurement, metabolism, mind, mole, molecule, motion, mutation, natural selection, nebula, neuron, organism, osmosis, particle, periodic table, pH, photosynthesis, planet, pollution, pressure, probability, protein, pulsar, quantum, quark, radioactivity, reaction, relativity, reproduction, research, rule, science, scientific method, solution, space, species, star, stem cell, symbiosis, systems, technology, temperature, theory, time, tissue, tumour, Universe, vacuum, virus, wave.

Scientists might all learn something from this list, representing as it does how our everyday work is perceived by a small sample of bright youngsters. Alongside several words that could just as well have been listed 100 years ago (such as acid/base, magnetism, mole, scientific method), I was struck by the number of terms hinging on ethical issues in medicine and biology (6%), the theory of evolution (5%) and clinical terminology (5%). The Internet too is up there among giant scientific words such as 'Universe' and 'atom'.

This thought-provoking collection of words suggests that, as working scientists, we need to care at least as much about science education as we do about publication.

Marco Prunotto Nephrology Unit Laboratory, Giannina Gaslini Children's Hospital, Largo G. Gaslini 5, 16147 Genova, Italy
e-mail: marco.prunotto@gmail.com

Animal research: too much faith in models clouds judgement

SIR — Your News Feature 'Standard model' (*Nature* **454,** 682–685; 2008) raises issues about the use of mouse models of disease that go well beyond the field of neurodegenerative disorders. As a former head of atherosclerosis research at GlaxoSmithKline laboratories, I can attest that the situation is similar for models of atherosclerosis and dyslipidaemia.

Informed users of mouse models are well aware of their limitations in relation to human pathology, so their expectations from drug studies and the relevance of these to humans are tailored accordingly. But it seems to me that, beyond this relatively small group of practitioners, wider concerns arise.

These might be viewed as a criticism of the rigour of much of

the dialogue between preclinical and clinical research. One is about overuse of the glib term "animal model of disease X", which raises expectations and clouds proper interrogation of experiments. As you suggest, it is better to consider a mouse model as primarily one of mechanism and to make a reasoned extrapolation to humans from there.

This approach links in more closely with the current critical preoccupation with translational research. Another concern is one that permeates all science, namely the tendency to regard the model (in whatever form it is expressed) as being identical to its prototype, often coupled with the idea of a 'complete explanation', which is, of course, illusory.

Keith Suckling Welwyn Garden City, Herts AL8 7NH, UK
e-mail: keith@suckling291.freeserve.co.uk

Animal research: raise standards to protect patients

SIR — Your News Feature 'Standard model' (*Nature* **454,** 682–685; 2008) highlights problems in using mice as models for human neurodegenerative disease. But these pitfalls may be widespread in animal research (see, for example, P. Perel *et al. Br. Med. J.* **334,** 197; 2007).

Animal studies are not always well designed and null results are rarely published. The standards should be raised to be comparable to those already established in clinical research. Preclinical research projects using animals should be prospectively registered and systematically reviewed. Also, they should be properly designed by using randomization, adequate sample sizes and blinding for evaluation of outcome.

Any rigorous scientific research requires investigators to set their basic hypotheses in the context of what is already published, to avoid unnecessary replication and to justify the new study. However, there is often a failure in animal research to apply these standards. This can expose research volunteers and patients to flawed and inadequate research and put biotech investors at risk of substantial long-term financial loss.

Supporting the introduction of these requirements for animal research would improve the body of evidence available to policy-makers and investors with regard to human health. It would also minimize wastage of laboratory animals and improve patient safety.

Susan Green SABRE Research UK, PO Box 18653, London NW3 4UJ, UK
e-mail: office@sabre.org.uk
Competing interest: SABRE Research UK is a charity that raises awareness of these issues in animal research.

'Which science book should the next US president read?' pages 464–467

Energy: efficiency gains alone won't reduce emissions

SIR — Your News Feature 'Electricity without carbon' (*Nature* **454,** 816–823; 2008) claims that the easiest way to cut the carbon released by electricity generation is to "increase efficiency". Although efficiency measures do have the potential to cut emissions, this applies only under the optimistic assumption that efficient energy converters will not be used more than their inefficient predecessors. This assumption has been challenged by Vaclav Smil in his Correspondence 'Long-range energy forecasts are no more than fairy tales' (*Nature* **453,** 154; 2008), who argues "We have known for nearly 150 years that, in the long run, efficiency gains translate into higher energy use".

Given the thermodynamic and economic limitations of efficiency measures, the easiest way of reducing emissions is also the simplest: cut usage of existing energy converters. For example, incentive-based regulation, real-time electricity metering and rethinking behaviour could all considerably reduce demand for electricity and therefore its associated emissions. In combination with efficiency gains and renewable energy sources, such an approach would create better conditions for a smooth transition to a carbon-neutral economy.

Robin Lovelace Norton Canon, Hereford HR4 7BP, UK
e-mail: Rob00x@gmail.com

Energy: time to consider heavy-metal nuclear coolants?

SIR — You provide an admirable summary of electricity generation without fossil fuels in your News Feature 'Electricity without carbon' (*Nature* **454,** 816–823; 2008). I would add to the section on nuclear fission the six concepts being studied by the US Department of Energy in its programme for considering the next generation of nuclear technologies.

One of these recommends using heavy-metal coolants, which avoids some of the disadvantages associated with water-cooled systems. Heavy-metal coolants are inherently safer and produce a much smaller fraction of waste for disposal (see, for example, E. P. Loewen *Am. Sci.* **92,** 522–531; 2004).

Donald E. Hirsch PO Box 196, West Boothbay Harbor, Maine 04575, USA
e-mail: dehirsch1924@yahoo.com

Big data: open-source format needed to aid wiki collaboration

SIR — Your News Feature 'Wikiomics' (*Nature* **455,** 22–25; 2008) points out that the growing online wiki-style community collaboration offers a smart way to respond to torrents of new data. Although lack of recognition and credit may prevent scientists from actively participating, this may be only part of the story. Reasons for not giving back to the community can be more complicated.

Handling and analysing big data sets is becoming more and more challenging. Wiki-style information is convenient, and the GenMAPP pathway database you mention is a good example. But the lack of a unified data format for facilitating the exchange from one database to another can be a killer for someone trying to submit data. There is currently no de facto standard on pathway-data format, which severely diminishes data portability.

The scientific community and commercial bodies should come up with a consortium to resolve this issue by creating an open-source format, instead of developing their own pathway-drawing tools. This would facilitate knowledge-sharing and allow more focused development.

The consortium could also serve as a hub to consolidate wiki pages. With the evolution of more and more similar wiki pages, information is at risk of becoming redundant and fragmented. The consortium might provide a forum for discussion among scientists from different disciplines to arrive at a consensus on this issue. This would also give the data-submitter more security and remove concerns about database discontinuity.

It all sounds easy, but in practice would be difficult. No one wants to waste what they have developed. However, with the cost of genomic research being reduced — take the 'US$1,000 genome' project, for example — a large number of big databases should materialize in the not-so-distant future. A more efficient wiki community is sorely needed to cope with this advancement.

Tin-Lap Lee Section on Developmental Genomics, Laboratory of Clinical Genomics, Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, Maryland 20892, USA
e-mail: leetl@mail.nih.gov

Big data: teaching must evolve to keep up with advances

SIR — The Editorial introducing your 'big data' special issue ('Community cleverness required' *Nature* **455,** 1; 2008) highlights the entrenched "institutions and culture of science" as barriers to "taking full advantage of electronic data". Science education is part of that entrenchment, so more thought should be invested in the tight relationship between the advancement of science and our systems for educating future scientists.

The growth of web-based, collaboratively developed and publicly accessible research data provides an important opportunity to transform the ways in which students are taught science. Failure to challenge the status quo in science education will lead to an even greater divergence between scientific practice and education, which could potentially undermine the continued rapid evolution of scientific research. If the use of cyberinfrastructure is going to be "as integral to the practice of science as publishing and reading papers", as Lincoln Stein has suggested (*Nature Rev. Genet.* **9,** 678–688; 2008), then we must invest in developing the human infrastructure necessary to realize these new approaches to doing science .

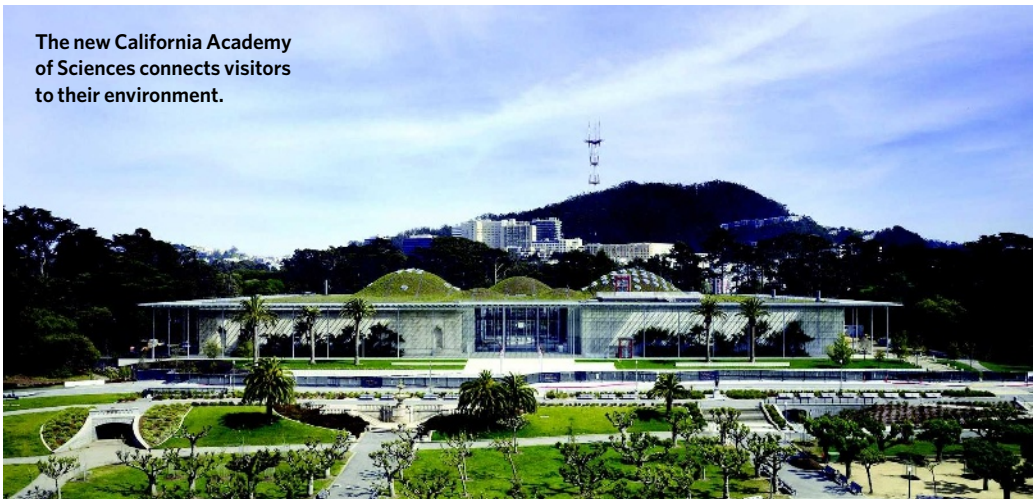
The opportunities to engage undergraduates in realistic scientific-research experiences using 'big data' and other e-science resources are as diverse as the research itself (see, for example, the BioQUEST Curriculum Consortium site http://bioquest.org). Unfortunately, computational infrastructure and successful model projects are still not adequate to ensure that the next generation of scientists will be prepared to work in a rapidly evolving research environment.

Most undergraduate science educators completed their graduate training more than 10 years ago and so have limited experience in working with modern e-science resources. Yet today's students are entering a scientific workforce in which they are expected to have skills in areas such as data mining, modelling, visualization and annotation.

Given the growing disparity between the rapidly evolving world of research and an entrenched culture of science education, the future of science depends less on "cleverness" than on our commitment to preparing future scientists to work with 'big data'.

Samuel Donovan Department of Biological Sciences and BioQUEST Curriculum Consortium, 4249 Fifth Avenue, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA
e-mail: sdonovan@pitt.edu

Contributions may be sent to correspondence@nature.com.



The new California Academy of Sciences connects visitors to their environment.

T. GRIFFITH

A sustainable home for science

The California Academy of Sciences reopens this week after a US$488-million redesign. Inspired by its setting in San Francisco's Golden Gate Park, the landmark museum harnesses sustainable building technology to kindle a sense of scientific wonder.

The original academy was founded in 1916. Severe structural damage caused by the 1989 Loma Prieta earthquake prompted executives to embark on the largest museum-rebuilding project in the world for a century. Architect Renzo Piano, 1998 winner of the Pritzker Architecture Prize, drew up plans for a new sustainable home for the academy, declaring that "in the twenty-first century, the most inspiring element for architecture is the fragility of Earth".

The academy will house 400 employees — including 100 working scientists — and tens of thousands of living organisms, from mangrove trees to African penguins, South American ants and tropical corals of the Philippines. Piano's design showcases this biodiversity while incorporating features to save energy, water and materials. For instance, the building uses reclaimed steel, and 68% of the insulation is made from recycled blue denim jeans.

The museum's signature feature is its living roof. Boasting seven undulating hills and a meadow of colourful native Californian wildflowers, the roof combines aesthetic and engineering functions. The seven hills reference the peaks encompassed by San Francisco's original city limits, and the wildflowers that have been chosen will bloom all year round. This hectare of plants will stop 7 million litres of rainwater run-off from flowing to the ocean each year, and the 20-centimetre-thick mat of greenery helps to insulate the building. The landscaped slopes funnel cool air into the interior courtyard, and automated skylights open and shut to reduce the need for electric lighting, heating and cooling systems. A canopy surrounding the roof's edge offers shade and supports solar panels to further cut energy costs.

Natural light and air pour through the exotic exhibits: a 30-metre-tall living rainforest with live insects, birds and bats, a remodelled planetarium, and the deepest living coral reef display constructed anywhere in the world. Piano fulfils his goal of connecting the museum's visitors with the environment outside. "The chance to be in the centre of Golden Gate Park is immense, so you have to be transparent," he explains.

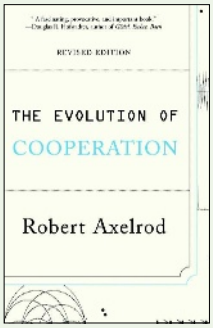
The bold new academy declares that we should rethink science's position in society. No longer secreting their wisdom in an ivory tower, scientists must today interact openly with the world. Piano shows that this move to transparency can be inspiring rather than frightening, and that science itself stands to benefit. ■

Erika Check Hayden is a correspondent for *Nature* based in San Francisco.

California Academy of Sciences
San Francisco, California
Opens 27 September 2008

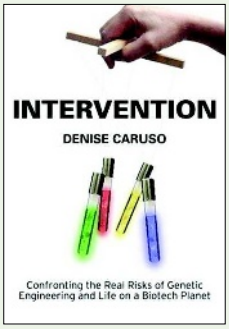
The Evolution of Cooperation by Robert Axelrod
(Basic Books, 1984)

The next US leader should be concerned by the instability of intelligent life, so Al Gore's *An Inconvenient Truth* (Rodale, 2006) is a good choice. Dinosaurs hung on for some 160 million years, but how long will humans survive? We are causing such irreversible ecosystem destruction that we will eliminate our own habitat. Technological advances alone cannot fix the problem. To reach a solution, humans must cooperate on a global scale, requiring us to show wisdom, generosity and respect. A classic from which we may all learn is Robert Axelrod's book. Martin Nowak, professor of biology and mathematics at Harvard University.



Intervention by Denise Caruso
(Hybrid Vigor Press, 2006)

In these uncertain times, the presidential incumbent might be less well served by all the good news about science. As a cure for complacency, recall that the mathematics of finance has been corrupted by being put at the service of fantasy and greed in the economic crisis that is now gripping the United States. Mathematics has enabled assets to be given fictitious valuations that even now threaten catastrophe. Scientific research in the service of industry and the state is not immune to such pressures. A cautionary tale is told by Denise Caruso in *Intervention: Confronting the Real Risks of*

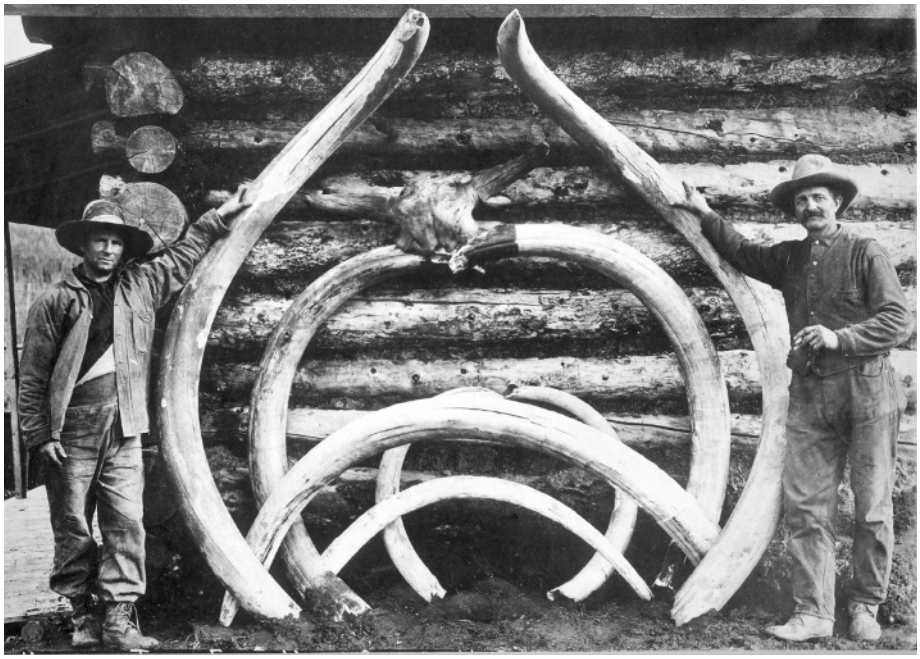


BOOKS & ARTS

Land of giants

Expansion of the railways across the western United States changed the face of vertebrate palaeontology, and perhaps the country itself, explains **Ross MacPhee**.

The Legacy of the Mastodon: The Golden Age of Fossils in America
by Keith Thomson
Yale University Press: 2008. 386 pp.
$35, £22.50



US LIBRARY OF CONGRESS

Huge mastodon bones from the US midwest dispelled the myth that all New World mammals were puny.

Vertebrate fossils are rarely found in nature, but prominent exceptions are in the rich interior geological basins of western North America. The arrival of the railways in these desolate regions in the mid-nineteenth century set the stage for a 'bone rush', as Keith Thomson describes in his new book, *The Legacy of the Mastodon*. The colourful characters who forged their reputations as palaeontologists in this treasure trove, Thomson argues, contributed to the evolving image of the United States as a land of giants, figuratively as well as literally.

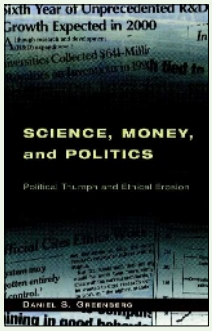
American vertebrate palaeontology got off to a slow start. Hardly anyone living in the United States at the beginning of the nineteenth century considered collecting old bones to be worthwhile, apart from a few savants associated with the American Philosophical Society, founded in 1745, and the Academy of Natural Sciences, created in 1812. Another notable exception was the third US president Thomas Jefferson, who held office between 1801 and 1809. When Meriwether Lewis and William Clark set out on their 1804 cross-country expedition to the Pacific, Jefferson asked them to look out for the animal owner of the 'great claw', a huge distal phalanx that had been found at the delightfully named Big Bone Lick in Kentucky. Jefferson believed that the claw belonged to a lion-like creature that still lurked in the west. In fact, it belonged to a ground sloth, later named *Megalonyx*. Jefferson, who did not think that species extinction was part of nature's plan, might have been disappointed to learn that *Megalonyx* had died out 10,000 years previously.

Public as well as scholarly interest was occasionally sparked when the remains of vanished beasts turned up under the plough or in the roadbed, especially if they could be portrayed as hideous devourers of flesh removed from creation by a merciful God. In 1801, the Peale family of painters and showmen improved gate takings at their museum by inverting the tusks on a mastodon skeleton they had pieced together, creating the impression that it had been a carnivore armed with huge canines rather than a primitive kind of elephant.

Thomson explains that sideshow became science only after 1860, thanks to the opening of the west by the US railways. Ease of travel

Which science book should the next US president read?

Science, Money, and Politics by Daniel S. Greenberg
(Univ. Chicago Press, 2001)



What a president needs to understand is not science — which science, after all? — but the role of scientific expertise in the democratic political process. Daniel S. Greenberg is the outstanding writer on the politics of modern US science, and this is his most pertinent book. He takes for granted that the institutional framework for mobilizing and channelling scientific expertise works fairly well and should not be subject to cynical subversion. But he also understands that scientists enter a political arena when they advise, and presidents must take decisions that reflect social and political priorities. Scientists, like anyone else, must hope that presidents have good priorities. No book of biology, physics or meteorology will ensure that they do.
Steven Shapin, professor of the history of science, Harvard University.

The Blind Watchmaker by Richard Dawkins
(W. W. Norton, 2006; first published by Longman, 1986)
There is a crisis in scientific literacy in the United States: only 25% of Americans accept our evolution from ape-like ancestors, yet 74% believe in angels. Republicans make it worse by proposing that creationism should be taught alongside evolution in public-school science classes. Anyone aspiring to be president should have a basic acquaintance with evolution

created new opportunities for agriculture, white settlement and political domination in the middle third of the country, and for the development of geology and vertebrate palaeontology. Fossil collecting had until then been an accidental affair. Now, targeted expeditions to fossil-rich areas such as the Bridger, Green River and Wasatch basins became feasible.

Among the first to take advantage were Joseph Leidy and Ferdinand V. Hayden. They were soon superseded by the two dark personalities of US palaeontology in the 1870s and 1880s, Edward Drinker Cope and Othniel Charles Marsh. Apart from their scientific achievements, these men are remembered mainly for their lengthy, unpleasant and demeaning 'bone wars', in which they vied to outdo each other by any means possible in publishing new taxa.

Thomson covers the usual ground but provides new insights into the men that Cope and Marsh hired to do most of the collecting. These exploited and mistreated workers — ex-cowboys, clerks and roughnecks — were sent to hunt fossils in some of the harshest places in the country. They were expected to spy on rival teams, steal their specimens if possible, and generally act like boneyard bandits. Marsh's actions were particularly egregious. At the close of the 1879 season on Como Bluff in Wyoming, he instructed William Harlow Reed to destroy any unexcavated bones at one quarry so that others, namely Cope or those working for him, couldn't have them. Such actions today would have the perpetrator sent up before an ethics committee and thrown out of academia. But not back then: Marsh later became president of the National Academy of Sciences.

As sordid as this phase was in the development of US vertebrate palaeontology, the advances in knowledge were enormous. Thomson, himself a palaeontologist and evolutionary biologist, points out that when annual fieldwork became standard practice in the late nineteenth century, institutions such as the Smithsonian and the American Museum of Natural History were able to amass the huge specimen-based collections for which they are famous. Better methods for fossil discovery and preservation, such as sieving and plaster jacketing, were rapidly developed, thanks much more to the ingenuity of poorly paid collectors such as John Bell Hatcher than to the quarrelling professors in the east. And hundreds of new species were added to lists of the extinct fauna of North America. Some of these proved useful for the geological correlation of isolated basins, a practical benefit at the time. By 1890, when Thomson closes his narrative, vertebrate palaeontology had changed from a minor diversion practised by few to a nascent professional discipline, mirroring the transformation of the United States from a marginal power at the century's beginning into the political and economic colossus it had become by the end.

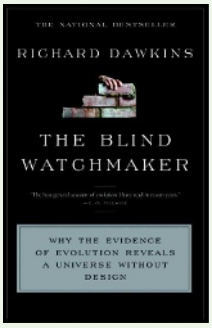
"It is hard to imagine how the harmless activity of fossilizing could have contributed to nation building."

Thomson makes much of this parallel progression, arguing that conditions in the early United States produced a peculiarly American style of palaeontology. The result is a view of history in which past events are interpreted as if they led inexorably to some predetermined end. One provocative argument in the book is that vertebrate palaeontology was destined to "play a role in the development of the myths of American nationhood". The only myth examined in detail is Comte de Buffon's 1766 theory of New World degeneracy. This held that the animal and human denizens of the Americas were physical and moral weaklings compared with their Old World counterparts. One of Buffon's 'proofs' — the lack of large, powerful mammals in North and South America — evaporated with the discovery of giant beasts such as the mastodon, much to Jefferson's satisfaction. Rather than myths as such, Thomson is more concerned with beliefs about opportunity and image. Where else but in the western reaches of the expanding United States could one find an unexplored vastness, redolent with the possibility of great discoveries and great success? Certainly, nineteenth-century North America produced fossils that had never before been seen. Nonetheless, it is hard to imagine how the relatively harmless activity of fossilizing could have contributed to nation building, then or now. Jefferson aside, Thomson's other examples are unconvincing. It seems improbable that the feckless Peale family really believed they were displaying "an aggressive symbol of American power" when they exhibited their reassembled mastodon skeleton in England in 1802.

For the palaeontology buff, Thomson supplies some unusual and amusing titbits, including some politically inspired doggerel that assails Jefferson's interest in both Pleistocene proboscideans and his mistress Sally Hemings. He also describes the machinations behind the first effort in 1858 to mount a dinosaur skeleton in a natural pose — a hadrosaur at the Academy of Natural Sciences. In appendices he introduces readers to geological time, and describes how Cope, much more than Marsh, contributed to the development of evolutionary theory. *The Legacy of the Mastodon* is a delicious read, instructive and amusing, and will entertain anyone who has wondered how we came to know the mastodon and its tribe. ■

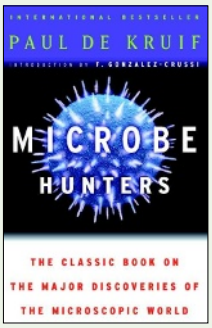
Ross MacPhee is curator of vertebrate zoology at the American Museum of Natural History, New York, New York 10024, USA. His forthcoming book on extinctions is *So Short a Season*.
e-mail: macphee@amnh.org

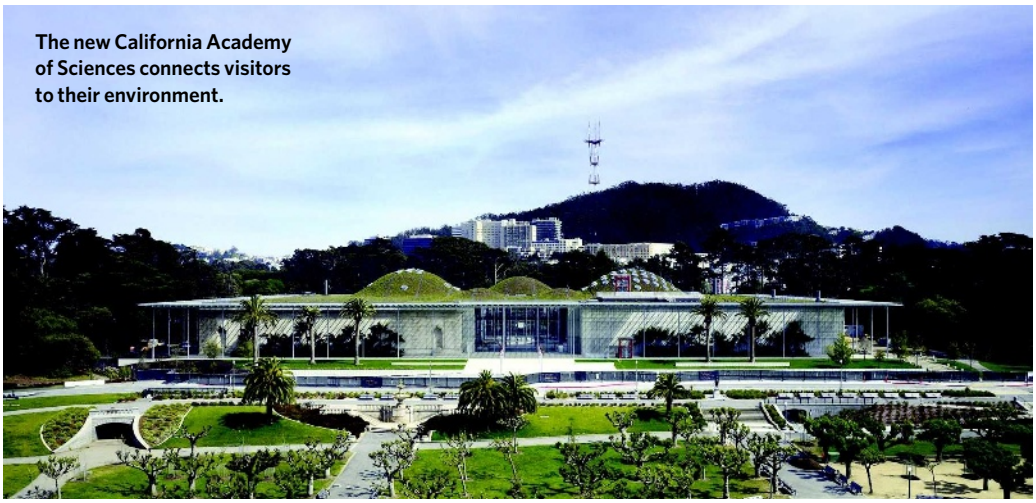
and with the masses of evidence that it's not just a theory, but a fact. Charles Darwin's *On the Origin of Species* comes to mind, but it is outdated and written in turgid Victorian prose that is uncongenial to modern readers. Future US leaders should read a short, popular work that lays out the evidence for evolution and dispels the spectres of creationism and intelligent design without dwelling on religion. Sadly, no book fills this niche. My attempt, *Why Evolution is True* (Viking, 2009), will be published only after the election. Until then, I suggest Richard Dawkins's brilliant exposition of natural selection. If a presidential candidate doesn't accept evolution after reading this book, there is no hope.
Jerry Coyne, professor of ecology and evolution, University of Chicago, Illinois.



Microbe Hunters by Paul de Kruif
(Houghton Mifflin Harcourt, 2002; first published by Harcourt, 1939)

Without question, the next president should read *Microbe Hunters* by Paul de Kruif. Probably more bioscientists and physicians have been stimulated to adopt their careers after reading this captivating book than any other. It is easy to read but still relevant, and might help a president to understand the life sciences and the commitment of life scientists to their work.
Rita Colwell, chairman of Canon US Life Sciences and distinguished university professor at the University of Maryland, College Park, and at Johns Hopkins University Bloomberg School of Public Health, Maryland.





The new California Academy of Sciences connects visitors to their environment.

T. GRIFFITH

A sustainable home for science

The California Academy of Sciences reopens this week after a US$488-million redesign. Inspired by its setting in San Francisco's Golden Gate Park, the landmark museum harnesses sustainable building technology to kindle a sense of scientific wonder.

The original academy was founded in 1916. Severe structural damage caused by the 1989 Loma Prieta earthquake prompted executives to embark on the largest museum-rebuilding project in the world for a century. Architect Renzo Piano, 1998 winner of the Pritzker Architecture Prize, drew up plans for a new sustainable home for the academy, declaring that "in the twenty-first century, the most inspiring element for architecture is the fragility of Earth".

The academy will house 400 employees — including 100 working scientists — and tens of thousands of living organisms, from mangrove trees to African penguins, South American ants and tropical corals of the Philippines. Piano's design showcases this biodiversity while incorporating features to save energy, water and materials. For instance, the building uses reclaimed steel, and 68% of the insulation is made from recycled blue denim jeans.

The museum's signature feature is its living roof. Boasting seven undulating hills and a meadow of colourful native Californian wildflowers, the roof combines aesthetic and engineering functions. The seven hills reference the peaks encompassed by San Francisco's original city limits, and the wildflowers that have been chosen will bloom all year round. This hectare of plants will stop 7 million litres of rainwater run-off from flowing to the ocean each year, and the 20-centimetre-thick mat of greenery helps to insulate the building. The landscaped slopes funnel cool air into the interior courtyard, and automated skylights open and shut to reduce the need for electric lighting, heating and cooling systems. A canopy surrounding the roof's edge offers shade and supports solar panels to further cut energy costs.

Natural light and air pour through the exotic exhibits: a 30-metre-tall living rainforest with live insects, birds and bats, a remodelled planetarium, and the deepest living coral reef display constructed anywhere in the world. Piano fulfils his goal of connecting the museum's visitors with the environment outside. "The chance to be in the centre of Golden Gate Park is immense, so you have to be transparent," he explains.

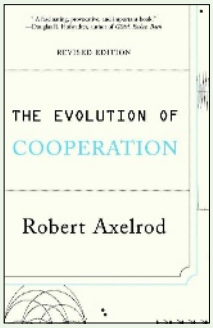
The bold new academy declares that we should rethink science's position in society. No longer secreting their wisdom in an ivory tower, scientists must today interact openly with the world. Piano shows that this move to transparency can be inspiring rather than frightening, and that science itself stands to benefit. ■

Erika Check Hayden is a correspondent for *Nature* based in San Francisco.

California Academy of Sciences
San Francisco, California
Opens 27 September 2008

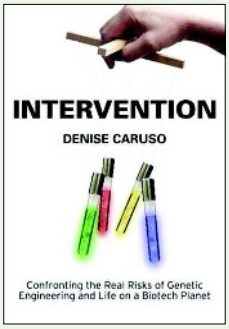
The Evolution of Cooperation by Robert Axelrod
(Basic Books, 1984)

The next US leader should be concerned by the instability of intelligent life, so Al Gore's *An Inconvenient Truth* (Rodale, 2006) is a good choice. Dinosaurs hung on for some 160 million years, but how long will humans survive? We are causing such irreversible ecosystem destruction that we will eliminate our own habitat. Technological advances alone cannot fix the problem. To reach a solution, humans must cooperate on a global scale, requiring us to show wisdom, generosity and respect. A classic from which we may all learn is Robert Axelrod's book. Martin Nowak, professor of biology and mathematics at Harvard University.



Intervention by Denise Caruso
(Hybrid Vigor Press, 2006)

In these uncertain times, the presidential incumbent might be less well served by all the good news about science. As a cure for complacency, recall that the mathematics of finance has been corrupted by being put at the service of fantasy and greed in the economic crisis that is now gripping the United States. Mathematics has enabled assets to be given fictitious valuations that even now threaten catastrophe. Scientific research in the service of industry and the state is not immune to such pressures. A cautionary tale is told by Denise Caruso in *Intervention: Confronting the Real Risks of*



Saving public universities

Unmaking the Public University: The Forty-Year Assault on the Middle Class
by Christopher Newfield
Harvard University Press: 2008. 408 pp. $29.95, £19.95, €22.50

Higher education in the United States is shared between public and private institutions. The former, such as the University of Michigan or the University of California, receive annual funding from their states. The latter, such as Stanford University or Harvard, rely on private revenue sources, especially student fees and income from endowments. Both types compete for research grants from the federal and state governments. The public universities were founded and are funded by their respective states to promote the public interest, especially agricultural, economic and social development.

In his book *Unmaking the Public University*, Christopher Newfield, professor of English at the University of California in Santa Barbara, draws attention to the drop in funding for public universities since the 1970s. Its decline, he argues contentiously, is a result of a campaign by "cultural conservatives" who brand these institutions as breeding grounds for a 'liberal' middle class.

Newfield's latest work succeeds his 2003 book *Ivy and Industry* (Duke University Press), which explored the problematic relationship between business and US universities. He argued that the liberal-arts tradition and capitalist culture are contradictory forces that create conflicts for both the academy and the students who go on to comprise much of the middle class. He challenged the modern North American university to promote the humanities while drawing on the benefits of business organization and dispensing with its negative aspects.

Now, Newfield builds on his earlier work by presenting a complex, historical narrative of a new middle class. The new cohort is composed of multiracial, progressive and upwardly mobile graduates of public universities. They are the knowledge workers in the knowledge economy. The public universities taught them the collaborative working styles, research skills and technical abilities that enabled the success of cutting-edge companies such as those in Silicon Valley, California. Newfield suggests that cultural conservatives targeted this group because they threatened the power of the traditional business and political elites.

Newfield identifies the major adversaries of this new middle class and the public university as the same conservative culture warriors who attacked the liberal influence on campus. His most striking claim is that for the latter, "The ultimate prize was the reduced cost and status of the middle class that the public university created". Newfield describes how conservatives, in league with powerful figures such as Ronald Reagan, a governor of California and later US president, influenced the reduction of funding for public universities, crippling this new middle class. The author also blames universities for compromising their institutional independence through their business dealings with government and corporations.

Newfield suggests some obvious remedies. Racial equality must be re-established as a primary national goal; the public universities must broaden their access and raise their academic quality; funding must be increased; and human development must be promoted together with economic opportunity. Yet he neglects key issues that are familiar to those engaged in the current debates over public higher education, making his thesis questionable.

A more plausible reason for the decreasing financial support for public universities, for example, is the perennial battle for scarce resources in state-government budgets. Public higher education must compete for funds with escalating health-care costs, increasing primary- and secondary-school needs, and rising demands by other government agencies. Thus, public universities have responded by supplementing their state funds through additional fund-raising, joint projects with private companies and higher student fees.

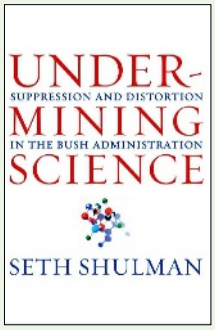
Newfield gives only one passing reference to the important and growing phenomenon of community colleges, which offer two-year academic and vocational degrees to students who cannot or choose not to attend a four-year course at a college or university. Nationally, there are approximately 6.7 million community-college students, and a significant number move on to further study on four-year degree programmes. Given that community colleges serve as a major gateway to a senior college education, this is a critical omission in the book.

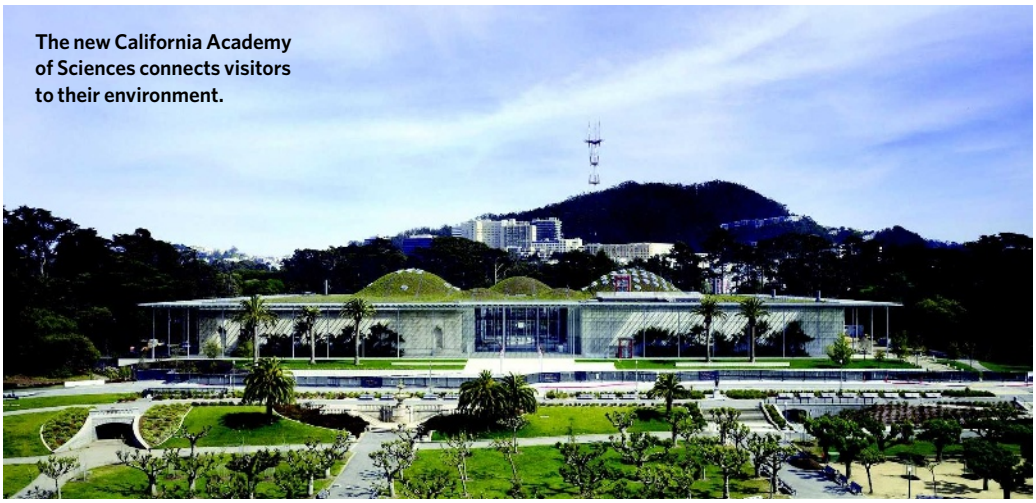
Another weakness is the lack of historical background. Unlike privately funded institutions, state funding has made economic development an integral part of the public university's mission for their home state. Prominent in this regard are the 'land grant' universities created by the Morrill Acts of 1862 and 1890, which were required to emphasize instruction in the agricultural and mechanical arts and sciences.

One could get the impression from reading *Unmaking the Public University* that the cultural battle has been one-sided. The book makes no mention of the major national higher-educational organizations such as the American Association of State Colleges and Universities, the National Association of

***Genetic Engineering and Life on a Biotech Planet*. Her story is one-sided, but it is well researched and she recommends a feasible system of risk analysis. The book will alarm some and anger others, but the debate is urgently needed.**
Jerry Ravetz, fellow of the James Martin Institute, University of Oxford, UK.

Undermining Science by Seth Shulman
(Univ. California Press, 2006)
Seth Shulman argues that George W. Bush's science policies recall those of Trofim Lysenko under Josef Stalin. The details are not pretty, the reporting is thorough and the evidence has not been credibly contested. Republican candidate John McCain might read it to see the extent to which neoconservative infrastructure is already in place in US science agencies, making it easier for him to continue the same policies without seeming to. His record on some science issues has been good, but his recent opinions, from energy to creationism in schools, have been drifting towards those of Bush. Democrat candidate Barack Obama might use Shulman's book to discover which recent science-agency appointees passed the test of right-wing fealty rather than that of scientific objectivity, thereby suggesting where he might make replacements. But such a task will not be easy. Shulman reports that the present administration has so thoroughly sown loyalists of questionable competence into science bodies — from NASA to the US Weights and Measures division — that it will take a considerable effort to root them out.
Kevin Padian, professor of integrative biology and curator of the Museum of Paleontology, University of California, Berkeley.





The new California Academy of Sciences connects visitors to their environment.

T. GRIFFITH

A sustainable home for science

The California Academy of Sciences reopens this week after a US$488-million redesign. Inspired by its setting in San Francisco's Golden Gate Park, the landmark museum harnesses sustainable building technology to kindle a sense of scientific wonder.

The original academy was founded in 1916. Severe structural damage caused by the 1989 Loma Prieta earthquake prompted executives to embark on the largest museum-rebuilding project in the world for a century. Architect Renzo Piano, 1998 winner of the Pritzker Architecture Prize, drew up plans for a new sustainable home for the academy, declaring that "in the twenty-first century, the most inspiring element for architecture is the fragility of Earth".

The academy will house 400 employees — including 100 working scientists — and tens of thousands of living organisms, from mangrove trees to African penguins, South American ants and tropical corals of the Philippines. Piano's design showcases this biodiversity while incorporating features to save energy, water and materials. For instance, the building uses reclaimed steel, and 68% of the insulation is made from recycled blue denim jeans.

The museum's signature feature is its living roof. Boasting seven undulating hills and a meadow of colourful native Californian wildflowers, the roof combines aesthetic and engineering functions. The seven hills reference the peaks encompassed by San Francisco's original city limits, and the wildflowers that have been chosen will bloom all year round. This hectare of plants will stop 7 million litres of rainwater run-off from flowing to the ocean each year, and the 20-centimetre-thick mat of greenery helps to insulate the building. The landscaped slopes funnel cool air into the interior courtyard, and automated skylights open and shut to reduce the need for electric lighting, heating and cooling systems. A canopy surrounding the roof's edge offers shade and supports solar panels to further cut energy costs.

Natural light and air pour through the exotic exhibits: a 30-metre-tall living rainforest with live insects, birds and bats, a remodelled planetarium, and the deepest living coral reef display constructed anywhere in the world. Piano fulfils his goal of connecting the museum's visitors with the environment outside. "The chance to be in the centre of Golden Gate Park is immense, so you have to be transparent," he explains.

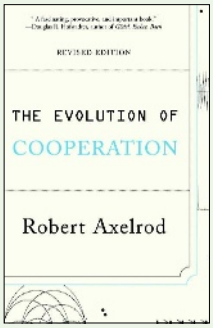
The bold new academy declares that we should rethink science's position in society. No longer secreting their wisdom in an ivory tower, scientists must today interact openly with the world. Piano shows that this move to transparency can be inspiring rather than frightening, and that science itself stands to benefit. ■

Erika Check Hayden is a correspondent for *Nature* based in San Francisco.

California Academy of Sciences
San Francisco, California
Opens 27 September 2008

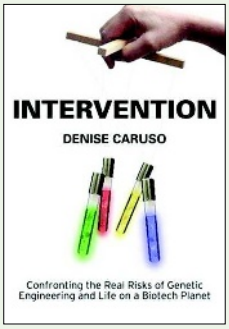
The Evolution of Cooperation by Robert Axelrod
(Basic Books, 1984)

The next US leader should be concerned by the instability of intelligent life, so Al Gore's *An Inconvenient Truth* (Rodale, 2006) is a good choice. Dinosaurs hung on for some 160 million years, but how long will humans survive? We are causing such irreversible ecosystem destruction that we will eliminate our own habitat. Technological advances alone cannot fix the problem. To reach a solution, humans must cooperate on a global scale, requiring us to show wisdom, generosity and respect. A classic from which we may all learn is Robert Axelrod's book. Martin Nowak, professor of biology and mathematics at Harvard University.



Intervention by Denise Caruso
(Hybrid Vigor Press, 2006)

In these uncertain times, the presidential incumbent might be less well served by all the good news about science. As a cure for complacency, recall that the mathematics of finance has been corrupted by being put at the service of fantasy and greed in the economic crisis that is now gripping the United States. Mathematics has enabled assets to be given fictitious valuations that even now threaten catastrophe. Scientific research in the service of industry and the state is not immune to such pressures. A cautionary tale is told by Denise Caruso in *Intervention: Confronting the Real Risks of*



Saving public universities

Unmaking the Public University: The Forty-Year Assault on the Middle Class
by Christopher Newfield
Harvard University Press: 2008. 408 pp. $29.95, £19.95, €22.50

Higher education in the United States is shared between public and private institutions. The former, such as the University of Michigan or the University of California, receive annual funding from their states. The latter, such as Stanford University or Harvard, rely on private revenue sources, especially student fees and income from endowments. Both types compete for research grants from the federal and state governments. The public universities were founded and are funded by their respective states to promote the public interest, especially agricultural, economic and social development.

In his book *Unmaking the Public University*, Christopher Newfield, professor of English at the University of California in Santa Barbara, draws attention to the drop in funding for public universities since the 1970s. Its decline, he argues contentiously, is a result of a campaign by "cultural conservatives" who brand these institutions as breeding grounds for a 'liberal' middle class.

Newfield's latest work succeeds his 2003 book *Ivy and Industry* (Duke University Press), which explored the problematic relationship between business and US universities. He argued that the liberal-arts tradition and capitalist culture are contradictory forces that create conflicts for both the academy and the students who go on to comprise much of the middle class. He challenged the modern North American university to promote the humanities while drawing on the benefits of business organization and dispensing with its negative aspects.

Now, Newfield builds on his earlier work by presenting a complex, historical narrative of a new middle class. The new cohort is composed of multiracial, progressive and upwardly mobile graduates of public universities. They are the knowledge workers in the knowledge economy. The public universities taught them the collaborative working styles, research skills and technical abilities that enabled the success of cutting-edge companies such as those in Silicon Valley, California. Newfield suggests that cultural conservatives targeted this group because they threatened the power of the traditional business and political elites.

Newfield identifies the major adversaries of this new middle class and the public university as the same conservative culture warriors who attacked the liberal influence on campus. His most striking claim is that for the latter, "The ultimate prize was the reduced cost and status of the middle class that the public university created". Newfield describes how conservatives, in league with powerful figures such as Ronald Reagan, a governor of California and later US president, influenced the reduction of funding for public universities, crippling this new middle class. The author also blames universities for compromising their institutional independence through their business dealings with government and corporations.

Newfield suggests some obvious remedies. Racial equality must be re-established as a primary national goal; the public universities must broaden their access and raise their academic quality; funding must be increased; and human development must be promoted together with economic opportunity. Yet he neglects key issues that are familiar to those engaged in the current debates over public higher education, making his thesis questionable.

A more plausible reason for the decreasing financial support for public universities, for example, is the perennial battle for scarce resources in state-government budgets. Public higher education must compete for funds with escalating health-care costs, increasing primary- and secondary-school needs, and rising demands by other government agencies. Thus, public universities have responded by supplementing their state funds through additional fund-raising, joint projects with private companies and higher student fees.

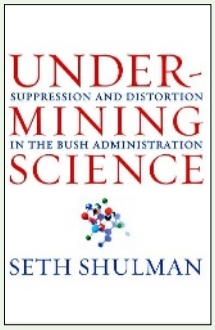
Newfield gives only one passing reference to the important and growing phenomenon of community colleges, which offer two-year academic and vocational degrees to students who cannot or choose not to attend a four-year course at a college or university. Nationally, there are approximately 6.7 million community-college students, and a significant number move on to further study on four-year degree programmes. Given that community colleges serve as a major gateway to a senior college education, this is a critical omission in the book.

Another weakness is the lack of historical background. Unlike privately funded institutions, state funding has made economic development an integral part of the public university's mission for their home state. Prominent in this regard are the 'land grant' universities created by the Morrill Acts of 1862 and 1890, which were required to emphasize instruction in the agricultural and mechanical arts and sciences.

One could get the impression from reading *Unmaking the Public University* that the cultural battle has been one-sided. The book makes no mention of the major national higher-educational organizations such as the American Association of State Colleges and Universities, the National Association of

***Genetic Engineering and Life on a Biotech Planet*. Her story is one-sided, but it is well researched and she recommends a feasible system of risk analysis. The book will alarm some and anger others, but the debate is urgently needed.**
Jerry Ravetz, fellow of the James Martin Institute, University of Oxford, UK.

Undermining Science by Seth Shulman
(Univ. California Press, 2006)
Seth Shulman argues that George W. Bush's science policies recall those of Trofim Lysenko under Josef Stalin. The details are not pretty, the reporting is thorough and the evidence has not been credibly contested. Republican candidate John McCain might read it to see the extent to which neoconservative infrastructure is already in place in US science agencies, making it easier for him to continue the same policies without seeming to. His record on some science issues has been good, but his recent opinions, from energy to creationism in schools, have been drifting towards those of Bush. Democrat candidate Barack Obama might use Shulman's book to discover which recent science-agency appointees passed the test of right-wing fealty rather than that of scientific objectivity, thereby suggesting where he might make replacements. But such a task will not be easy. Shulman reports that the present administration has so thoroughly sown loyalists of questionable competence into science bodies — from NASA to the US Weights and Measures division — that it will take a considerable effort to root them out.
Kevin Padian, professor of integrative biology and curator of the Museum of Paleontology, University of California, Berkeley.



State Universities and Land-Grant Colleges, or the American Council on Education, which have led the fight for the funding of public higher education at local, state and national levels.

Has the public university been unmade and Newfield's new middle class damaged? Hardly. Government funding has decreased in public universities, but they still have a critical role in US society. In the prestigious Association of American Universities, composed of the 60 leading research universities in the nation, more than half are public. The middle class is suffering some economic reversals, but the reasons are complex. Public higher education will continue to have a leading role in providing access to a good-quality education at an affordable price, for those who wish to improve their standard of living and quality of life. ■

John B. Clark is interim chancellor at the State University of New York, State University Plaza, Albany, New York 12246, USA.
e-mail: john.clark@suny.edu

Q&A: Science sketched out

The annual UK Big Draw festival, a month of nationwide workshops and talks launched this weekend in London, teaches people how to 'see' through drawing. **Terry Rosenberg**, head of design at Goldsmiths, University of London, explains how gaining skills in life drawing can help scientists perceive the world and communicate their results.

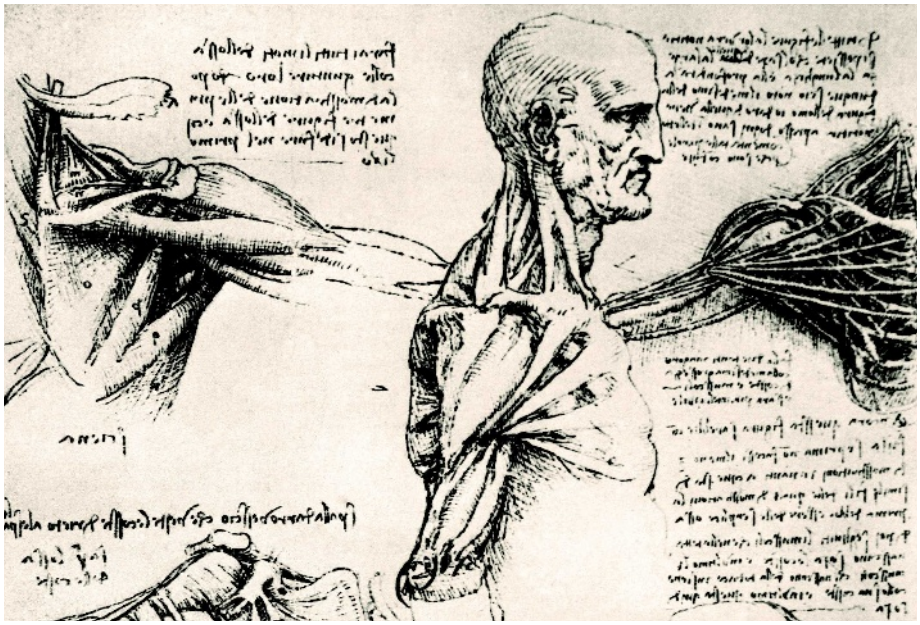
Drawing on Life
Wellcome Collection, London
26–28 September 2008
University College London
27 September 2008
www.thebigdraw.org

Why should scientists learn to draw?
Drawing can improve our understanding of perception by isolating different parts of the perceptual apparatus. When I first learned to draw, I would break up a scene into patches and try to reproduce each one. But this doesn't work — the eyes are constantly adjusting, the head moving, the light changing.

One teaching technique I use is to draw by moving the pencil in the same way that the eyes scan. This highlights the way the brain gathers information by making connections between different elements of the scene. Alberto Giacometti and other artists have done this more formally by plotting out such spatial relationships, into which a drawing can then be situated.

Leonardo Da Vinci studied anatomy using life drawing. Might students do so today?
Drawing moves from the general to the particular — looking at how the anatomy is articulated in practice; the positions, posture and movement of one bone over the other. I ask students to draw a posed skeleton, then draw a model in the same pose, building the muscles on top of the bones and resolving the tensions between them. You can then go back and infer what is inside the model.

Colouring and copying diagrams of the human body, as students do, is still drawing. They seem like strategies for rote learning, but when you look at an object and sketch it, you bring all your previous experience and knowledge to that moment.



BETTMANN/CORBIS

Life drawing is a useful learning method for anatomy students today — as it was for Leonardo da Vinci.

Can the act of drawing generate ideas?
Einstein claimed that his theories of relativity were merely hunches followed up by rigorous method. But how did he get the hunch? My design research looks at how laying out connections in time and space on the page can trigger new ideas. Rather than just seeking trends, our brains make nonlinear connections and homologies between juxtaposed objects, and original ideas result.

Might scientists use diagrams more effectively to communicate ideas?
Diagrams offer a quick way of making a hypothesis. Canonical representations such as hieroglyphs, where the symbol for a face might show it in profile, are often used. They communicate what is most important, but can ignore other elements of the underlying structure — the profile face hieroglyph misses out an eye, for example. And diagrams can also be driven too much by aesthetics: even anatomical diagrams use metaphors of the landscape, such as rivers, peaks and valleys. Scientists might learn from design research, which explores how information can be best conveyed spatially — as in the schematic map of the London Underground, for example.

Does drawing try to represent the 'truth'?
Subjective knowledge always intercedes in the processes that bind the hand to the eye. It has been said that drawing 'cleans things', but there is no ultimate image — every drawing tells you something different. ■

Interview by **Louise Whiteley**, a computational neuroscientist and writer based in London. She is deputy editor of *The Liberal* magazine.
e-mail: lewhiteley@gmail.com

State Universities and Land-Grant Colleges, or the American Council on Education, which have led the fight for the funding of public higher education at local, state and national levels.

Has the public university been unmade and Newfield's new middle class damaged? Hardly. Government funding has decreased in public universities, but they still have a critical role in US society. In the prestigious Association of American Universities, composed of the 60 leading research universities in the nation, more than half are public. The middle class is suffering some economic reversals, but the reasons are complex. Public higher education will continue to have a leading role in providing access to a good-quality education at an affordable price, for those who wish to improve their standard of living and quality of life. ■

John B. Clark is interim chancellor at the State University of New York, State University Plaza, Albany, New York 12246, USA.
e-mail: john.clark@suny.edu

Q&A: Science sketched out

The annual UK Big Draw festival, a month of nationwide workshops and talks launched this weekend in London, teaches people how to 'see' through drawing. **Terry Rosenberg**, head of design at Goldsmiths, University of London, explains how gaining skills in life drawing can help scientists perceive the world and communicate their results.

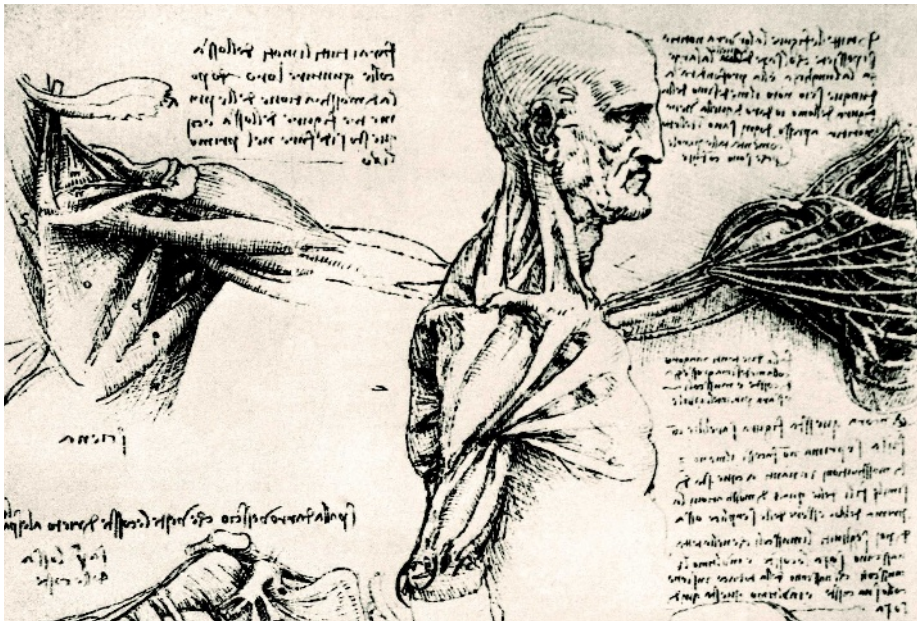
Drawing on Life
Wellcome Collection, London
26–28 September 2008
University College London
27 September 2008
www.thebigdraw.org

Why should scientists learn to draw?
Drawing can improve our understanding of perception by isolating different parts of the perceptual apparatus. When I first learned to draw, I would break up a scene into patches and try to reproduce each one. But this doesn't work — the eyes are constantly adjusting, the head moving, the light changing.

One teaching technique I use is to draw by moving the pencil in the same way that the eyes scan. This highlights the way the brain gathers information by making connections between different elements of the scene. Alberto Giacometti and other artists have done this more formally by plotting out such spatial relationships, into which a drawing can then be situated.

Leonardo Da Vinci studied anatomy using life drawing. Might students do so today?
Drawing moves from the general to the particular — looking at how the anatomy is articulated in practice; the positions, posture and movement of one bone over the other. I ask students to draw a posed skeleton, then draw a model in the same pose, building the muscles on top of the bones and resolving the tensions between them. You can then go back and infer what is inside the model.

Colouring and copying diagrams of the human body, as students do, is still drawing. They seem like strategies for rote learning, but when you look at an object and sketch it, you bring all your previous experience and knowledge to that moment.



BETTMANN/CORBIS

Life drawing is a useful learning method for anatomy students today — as it was for Leonardo da Vinci.

Can the act of drawing generate ideas?
Einstein claimed that his theories of relativity were merely hunches followed up by rigorous method. But how did he get the hunch? My design research looks at how laying out connections in time and space on the page can trigger new ideas. Rather than just seeking trends, our brains make nonlinear connections and homologies between juxtaposed objects, and original ideas result.

Might scientists use diagrams more effectively to communicate ideas?
Diagrams offer a quick way of making a hypothesis. Canonical representations such as hieroglyphs, where the symbol for a face might show it in profile, are often used. They communicate what is most important, but can ignore other elements of the underlying structure — the profile face hieroglyph misses out an eye, for example. And diagrams can also be driven too much by aesthetics: even anatomical diagrams use metaphors of the landscape, such as rivers, peaks and valleys. Scientists might learn from design research, which explores how information can be best conveyed spatially — as in the schematic map of the London Underground, for example.

Does drawing try to represent the 'truth'?
Subjective knowledge always intercedes in the processes that bind the hand to the eye. It has been said that drawing 'cleans things', but there is no ultimate image — every drawing tells you something different. ■

Interview by **Louise Whiteley**, a computational neuroscientist and writer based in London. She is deputy editor of *The Liberal* magazine.
e-mail: lewhiteley@gmail.com

ESSAY

D. PARKINS

Bellagio 1969: The green revolution

Agriculture in developing countries was transformed when scientists met aid officials and convinced them to invest in research. **Lowell S. Hardin** was there, and believes today's food crisis demands a similar vision.

In late April 1969, I took a taxi up a narrow, winding road outside Bellagio on Italy's Lake Como, to a meeting called by the president of the Rockefeller Foundation, George Harrar.

My destination was the foundation's conference centre at Villa Serbelloni, one of the most spectacular settings in the world. The conference had just one aim: to help solve the world food crisis. As we came in sight of the building, with its immaculately landscaped grounds set against snow-capped mountains, I knew I would remember the view, but had little inkling that the conference would have such a profound and lasting impact on world hunger.

The world food situation was becoming ever more precarious. Populations in developing countries were growing rapidly, and traditional farming systems were not producing enough to feed them. Two years before, the brothers William and Paul Paddock had published their best-seller *Famine — 1975! America's Decision: Who Will Survive?* (Little, Brown), predicting among other things that India would never feed its booming population. In both 1964 and 1965, the US Food for Aid Program had shipped 5 million tonnes of wheat aid to India. Many commentators thought the situation hopeless and famine inevitable.

At the Rockefeller and Ford foundations, we were more optimistic. We already had ten years' experience of collaborating with partners in developing countries in scientific efforts to accelerate food production. One reason that Harrar had called the Bellagio conference was to help agricultural aid organizations understand why and how science — and not food shipments — was a more sustainable way to deal with world hunger. Our thinking was that if the aid groups could grasp more fully the progress being made, they could mobilize the resources needed from governments and other donor organizations. Together, a sustained assault on world hunger could be made.

No time to waste

Thus, 24 of us met to thrash out a strategy for feeding the world's hungry. There were 16 leaders from the world's major foreign assistance agencies concerned with agricultural development — such as Adekke Boerma, director-general of the Food and Agriculture Organization of the United Nations, John Hannah, director of the US Agency for International Development, and Robert McNamara, president of the World Bank and former defence secretary to President John Kennedy — plus eight of us consultants from the science of food production. There were no women, even though women produce much of the developing world's food.

My colleague and mentor from the Ford

Foundation, Forrest 'Frosty' Hill, set the scene. In an ideal world, he said, developing countries would invest in education, research and infrastructure, building stronger public institutions and private businesses. These, in partnership with the better-educated farmers, would tackle the food crisis using science and technology, as had happened in industrialized countries during the agricultural revolution of the mid-twentieth century. "But," he said, "we are in a crisis. We cannot wait."

Urgent action was needed, but what? Many of the scientists favoured expanding the network of international agricultural research and training centres that was taking shape in developing nations. To illustrate why they thought this would be a good move, two scientists reported on the impact of the first two centres: one for improving wheat and maize, based in Mexico; and another centre that focused on rice, based in the Philippines.

The scientists were Sterling Wortman of the Rockefeller Foundation, whose idea the Bellagio conference was, and Robert Chandler, director-general of the International Rice Research Institute (IRRI) in the Philippines. Wortman described the work of the International Maize and Wheat Improvement Centre (CIMMYT) in Mexico. This was set up in 1966, although the strategy behind it dated from the pioneering Rockefeller–Mexico programme launched in 1943. At that time the country's average wheat yields were about 740 kilograms per hectare (11 bushels per acre). By 1967, average yields were 2,690 kilograms per hectare and Mexico was a net exporter of wheat. These remarkable results were achieved with new varieties that were stocky, disease-resistant, fast-growing and highly responsive to fertilizer and improved agricultural practices.

Success with rice came even quicker. Emulating the Mexican strategy, the IRRI had launched a plant-breeding programme in 1962. At that time, the average rice yield in south-east Asia was 1,500 kilograms per hectare. By 1965, the IRRI was developing semi-dwarf, stiff-strawed, disease-resistant and fertilizer-responsive plants that could double or triple yields. One of the new varieties had a potential yield of more than 9,000 kilograms per hectare. Farmers called them miracle plants, but they were the products of intensive research and testing, and generated through improved production practices and technologies.

Harrar and Hill had invented the 'centre model' for international agricultural research. The Ford and Rockefeller foundations had already developed and financed the IRRI and the CIMMYT, as well as the International Institute for Tropical Agriculture in Nigeria and the International Center for Tropical Agriculture in Colombia, both established in 1967. The capital and operating costs for this effort had reached US$41.3 million and more donors were needed if the model was to be expanded to include other countries, crops and animals. We desperately needed the aid organizations as major financial partners.

Gathering momentum

At one point, someone asked Hill if traditional farmers would adopt new technologies. It was the moment that the conference really began to gel. "Sure, if they are profitable enough," replied Frosty, in his homespun and persuasive tones. "In India, the new wheat varieties are in such demand that they have to guard their seed multiplication plots around the clock to prevent theft." Later the same day, when we described how investment in rice research had resulted in financial rates of return of more than 50%, McNamara stood up and said: "If you with your centres can generate returns like that I will help you raise the money you need." Hannah nodded, even suggesting that the US government might contribute a dollar for every three provided by other donors.

"Bellagio mobilized the world's agricultural-development organizations."

As this played out, I began to think: "Yes, we are getting to a meeting of minds." By the second day, it was clear the conference was going well. People from the aid side who had just met were on first-name terms. Equally important, aid people were talking to scientists. Serious conversations continued at tea breaks on the patio, during evening cocktails and at meals. By the closing session on the third day, our thinking was converging.

What did we agree on? First, that the key to increasing agricultural productivity in developing countries was to apply modern scientific techniques and technologies. Second, that setting up international centres of expertise in research and education was a proven shortcut to achieving this. Third, that the existing four centres should be fully funded and that another six to twelve centres should be created.

But there were concerns, too. We worried that a widespread green revolution could have unintended consequences, such as aggravating the inequalities between small farmers and large landowners. Furthermore, without careful management, intensified cropping could deplete soil and water resources and become unsustainable — which in some instances has happened. However, we concluded that world food needs outweighed such potential difficulties. Making advances in productivity sustainable remains a high-priority research goal today.

Bellagio was the catalyst. It mobilized the world's agricultural-development organizations to set in motion plans for rapidly increasing food production. It took the right mix of open-minded aid officials and dedicated scientists to achieve this, and it succeeded beyond any of our imaginings. After two follow-up conferences at Bellagio in the spring of 1970, it was agreed to set up a Consultative Group on International Agricultural Research, and in 1971 the organization that carries its acronym (CGIAR) was formed, under the leadership of the World Bank.

McGeorge Bundy, president of the Ford Foundation, called the CGIAR's creation "a remarkable chapter in the diplomacy of international assistance". Robert McNamara kept his pledge to help mobilize the necessary funding. By 1975, the year the Paddock brothers had forecast famine would strike India, the green revolution in Asia was under way. India, instead of starving, had achieved food independence. More than half the wheat and rice crops planted across Asia were high-yielding varieties.

Funding for the international centres rose substantially in the years after the Bellagio meeting. In 1969 the Rockefeller and Ford foundations provided $2.3 million for four centres. Today, the CGIAR has 64 donors providing more than $450 million a year in support of 15 centres and their 850 research scientists. Regretfully, some of the centres are critically underfunded, along with most of the agricultural research programmes with which they work.

As we confront today's food crisis, it is imperative that the green revolution in Asia is revitalized and a new one launched in Africa, which missed out almost entirely the first time round. As well as the research centres' ongoing work, promising new initiatives are under way, such as the Alliance for a Green Revolution in Africa, spearheaded by the Rockefeller and Bill & Melinda Gates foundations. Science and technology, including the exciting opportunities offered by genetically modified plants, have a lot more to offer the world's poor and hungry. As Norman Borlaug put it: "Responsible biotechnology is not the enemy; starvation is." ■

Lowell S. Hardin was a participant in the Bellagio Conference on Agricultural Development. He is assistant director of International Programs in Agriculture at Purdue University and was the Ford Foundation's program officer for Agriculture, Office of the Vice President, 1965–81.
e-mail: lhardin@purdue.edu

For more Meetings that Changed the World, see www.nature.com/nature/focus/meetings

NEWS & VIEWS

STRUCTURAL BIOLOGY

A moving story of receptors

Thue W. Schwartz and Wayne L. Hubbell

Animals sense light and chemical signals through proteins called G-protein-coupled receptors. The crystal structure of one such receptor in complex with a G-protein fragment shows how these receptors are activated.

The cell membranes of animals — everything from mammals to molluscs, insects and flatworms — contain proteins called G-protein-coupled receptors. Also known as 7TM receptors because of their seven transmembrane helices, they are involved in sensing light and a multitude of chemical signals: not only hormones and other transmitter molecules, but also odours, pheromones and flavours[1]. Many therapeutic drugs work by regulating the behaviour of specific 7TM receptors, and these proteins constitute the largest family of targets pursued by the pharmaceutical industry. Structural information about the receptors and their activation mechanisms is therefore clearly essential, but crystal structures have been available only for 7TM receptors in inactive forms. On page 497 of this issue, Scheerer *et al.*[2] report a long-awaited breakthrough: the first convincing structure of a 7TM receptor protein in an active conformation.

Until a year ago, the only high-resolution structures available for 7TM receptors were those for the inactive state of rhodopsin, a light receptor found in the rod cells of the retina[3]. Rhodopsin consists of a protein called opsin and a covalently bound ligand molecule called 11-*cis* retinal, which prevents the protein from signalling. Light converts the retinal to the active isomer (all-*trans* retinal), which stabilizes the active conformation of the receptor. As with all 7TM receptors, active rhodopsin binds and activates a subunit (the α-subunit) of an intracellular G protein (Fig. 1), and thereby translates the light signal into a cellular response. Several structures of rhodopsin that contain all-*trans* retinal have been obtained, but all of these essentially show the inactive conformation of the protein and offer few, if any, clues to the active state[3].

Rhodopsin has long been seen as the archetypal 7TM receptor, because particularly detailed biophysical and structural data were available for it. But last year saw the publication[4–6] of the first structures of a 7TM receptor — the β_2-adrenergic receptor — that is activated by a diffusible ligand, rather than light. More recently, the X-ray structure of a stabilized mutant of the β_1-adrenergic receptor was also solved[7]. These structures differed substantially from that of rhodopsin, mainly in their extracellular domains. In each case, however, the adrenergic receptors were crystallized with antagonists (receptor blockers), and so the structures determined are thought to be inactive conformations. Nevertheless, smaller, but still noteworthy, conformational differences in the intracellular parts of the transmembrane helices were observed compared with rhodopsin, especially for key amino-acid residues that are presumed to be involved in the activation process.

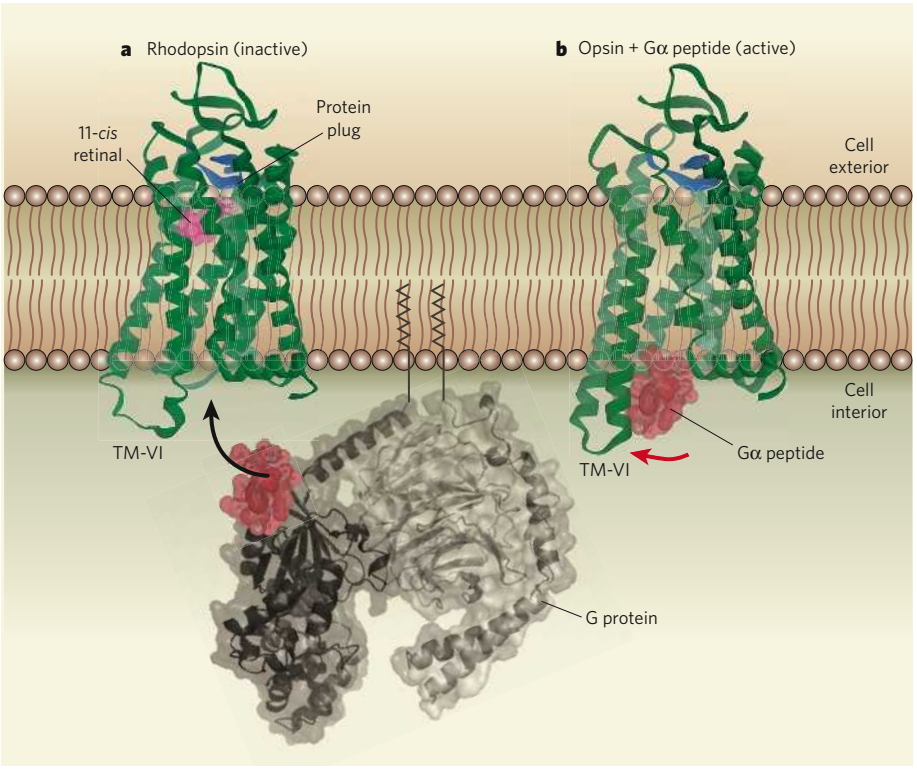


Figure 1 | Activation of a G-protein-coupled receptor. a, Rhodopsin, shown here in its inactivated conformation, is a light-sensing receptor found in cell membranes. It consists of a protein (opsin, green) and a ligand (retinal, pink, also shown in its inactivated conformation). When activated by light, rhodopsin binds to part of an adjacent G protein (binding region in red), triggering a cascade of biological responses. The protein plug (blue) is part of the extracellular domain of opsin, and immobilizes the extracellular transmembrane segments of the receptor. **b**, Scheerer *et al.*[2] have determined the activated structure of opsin in complex with the receptor-binding peptide fragment of the G protein (the Gα peptide). The most notable difference when compared with the inactivated receptor is that transmembrane helix 6 (TM-VI) has moved substantially outward (indicated by the red arrow), thereby creating the binding pocket for the G-protein peptide.

Scheerer *et al.*[2] now report the structure of opsin in the elusive activated state. In cells, the active conformation of a 7TM receptor is normally stabilized not only by an agonist (a receptor activator) at the extracellular side, but also by the α-subunit of the G protein at the intracellular side (Fig. 1). The latter effect was exploited by the authors, who crystallized opsin in complex with the main peptide fragment of the Gα protein to which opsin binds.

Compared with inactive receptors, the hallmarks of the activated structure are a relatively

large tilt of transmembrane helix 6 (TM-VI) and a smaller motion of helix 7 (TM-VII) at the inner surface of the cell membrane. These movements open a cleft that presents the ligand-binding site to the G protein. Scheerer and colleagues' results broadly match those of earlier biophysical studies[8], and agree quantitatively with those of work[9] that used pairs of 'spin labels' — which contain unpaired electrons that can be detected spectroscopically — to map out helix movement in activated rhodopsin. On the other hand, the authors' structure contrasts with a previously reported crystal structure[10] of activated rhodopsin (in the absence of a G protein) that differed little from the inactive state. Given the extensive biochemical and biophysical data in support of Scheerer and colleagues' structure for the activated state, it now seems likely that the previously reported structure[10] was modified by the conditions used to stabilize the crystals.

It might seem strange that the activating ligand, all-*trans* retinal, is not present in Scheerer and colleagues' opsin complex. But it is well established that active conformations of unliganded receptors, including opsin[11], exist in equilibrium with the inactive state. Indeed, 7TM receptors are often observed signalling with high activity through G proteins in the absence of agonist ligands. The crucial point in Scheerer and colleagues' structure[2] is that the observed active form is stabilized by a large excess of the G-protein fragment (the Gα peptide). It is the presence of this peptide that identifies the structure as a true active conformation. What is more surprising is that, in the absence of both all-*trans* retinal and the Gα peptide, opsin crystallizes in the same conformation as the authors report here (as reported earlier this year[12] by the same group).

The X-ray structure and orientation of the bound Gα peptide[2] are almost identical to those previously determined by nuclear magnetic resonance (NMR) studies[13] on solutions of a similar peptide interacting with activated rhodopsin. But Scheerer and colleagues' structure now shows how the peptide is stabilized by interactions with specific amino-acid residues that are exposed by the movement of TM-VI. Of particular note is that the peptide makes direct contact with the most evolutionarily conserved residue in 7TM receptors, an arginine residue at the intracellular end of TM-III.

The mechanism underlying the molecular recognition between opsin and the Gα peptide will probably be a good model for the interactions between 7TM receptors and the α-subunits of G proteins in general. The selectivity of these interactions is largely determined by the five carboxy-terminal residues of the α-subunits[14]. But it has been unexpectedly difficult to identify the corresponding selectivity-determining footprint on the receptors. The structure of Scheerer and colleagues' complex will therefore undoubtedly spur a wave of studies aimed at understanding this recognition process in detail. Moreover, interactions of 7TM receptors with G proteins, and their subsequent activation processes, are known to involve the recognition of additional structural elements on both the receptor and the G protein (Fig. 1). These processes can now be investigated in structure-based studies, as suggested by Scheerer and co-workers[2].

The observed conformation of activated opsin on the intracellular side of the cell membrane probably represents the activated state for 7TM receptors in general. But the relatively minor conformational changes that occur in the ligand-binding domain — which lies between the extracellular segments of the transmembrane helices — is likely to be unique to opsin. In rhodopsin, part of the extracellular region of opsin forms a protein 'plug' that fills the entrance to the main ligand-binding pocket, preventing movement of the transmembrane helices (Fig. 1). In contrast, in the previously determined structures of adrenergic receptors[5–7], the main ligand-binding crevice forms an open funnel towards the extracellular space. This allows ligands to diffuse in and out of the binding site, and could also allow inward movement of the extracellular segments of the transmembrane helices so that they surround the ligand. Such movement might aid ligand binding, and form part of a global 'toggle-switch' mechanism of receptor activation[8]. To test this possibility, we will need structures of 7TM receptors in complex not only with the G protein, but also with small-molecule agonists. ■

Thue W. Schwartz is at the Laboratory for Molecular Pharmacology, University of Copenhagen, Blegdamsvej 3, 2200 Copenhagen, Denmark. Wayne L. Hubbell is at the Jules Stein Eye Institute and the Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, California 90095, USA.
e-mails: tws@sund.ku.dk; hubbellw@jsei.ucla.edu

1. Pierce, K. L., Premont, R. T. & Lefkowitz, R. J. *Nature Rev. Mol. Cell Biol.* **3,** 639–650 (2002).
2. Scheerer, P. *et al. Nature* **455,** 497–502 (2008).
3. Schertler, G. F. *Curr. Opin. Struct. Biol.* **15,** 408–415 (2005).
4. Rasmussen, S. G. F. *et al. Nature* **450,** 383–387 (2007).
5. Rosenbaum, D. M. *et al. Science* **318,** 1266–1273 (2007).
6. Cherezov, V. *et al. Science* **318,** 1258–1265 (2007).
7. Warne, T. *et al. Nature* **454,** 486–491 (2008).
8. Schwartz, T. W., Frimurer, T. M., Holst, B., Rosenkilde, M. M. & Elling, C. E. *Annu. Rev. Pharmacol. Toxicol.* **46,** 481–519 (2006).
9. Altenbach, C., Kusnetzow, A. K., Ernst, O. P., Hofmann, K. P. & Hubbell, W. L. *Proc. Natl Acad. Sci. USA* **105,** 7439–7444 (2008).
10. Salom, D. *et al. Proc. Natl Acad. Sci. USA* **103,** 16123–16128 (2006).
11. Woodruff, M. L. *et al. Nature Genet.* **35,** 158–164 (2003).
12. Park, J. H. *et al. Nature* **454,** 183–187 (2008).
13. Koenig, B. W. *et al. J. Mol. Biol.* **322,** 441–461 (2002).
14. Conklin, B. R., Farzel, Z., Lustig, K. D., Julius, D. & Bourne, H. R. *Nature* **363,** 274–276 (1993).

Competing financial interests: declared (see online article for details).

SOLID-STATE PHYSICS

New order for magnetism

Eiji Saitoh

Physicists have come up with an innovative way of manipulating the direction of magnetization in a solid. The approach might be used to make low-power-consumption computer memory devices.

On page 515 of this issue, Chiba *et al.*[1] show how the direction of magnetization in a solid can be controlled electrically. The significance of this achievement lies in the fact that magnetic data-storage and memory devices, such as hard disks, depend on manipulating the direction of magnetization. But methods for doing this have so far been inefficient.

The fundamental property of a magnet is that, in a magnetic field, it aligns in the direction of the field (Fig. 1a). A magnet consists of atomic-scale magnets, called spins, and if a strong magnetic field is applied to a fixed magnet, the spins rotate to align with the field. As a result, the magnetization in the magnet — the average of the spin-magnetic moment — takes on the direction of the applied magnetic field (Fig. 1b). Magnetic disks exploit this principle. In a hard disk, for instance, information is stored on a disk-shaped magnet in the form of patterns of local magnetization and, to write information, a pulse of current is applied to a small electromagnet that scans the disk. But this method is indirect: a magnetic field exists between the current and the manipulated magnetization, a situation that wastes energy.

A promising alternative would be to apply a spin-polarized current directly to the magnet instead of using current to generate the magnetic field (Fig. 1c). Such a current would exert torque on the magnetization by exchanging spin-angular momentum with it as it passes through the magnet. Exploitation of this phenomenon, called spin torque[2], is expected to allow the development of compact magnetic memory devices that can run on low power consumption. The use of multiferroics[3], materials in which ferroelectricity and ferromagnetism coexist, has also been proposed. But direct control of the magnetization direction has not yet been achieved in such systems.

Chiba *et al.*[1] demonstrate a new way of manipulating the direction of magnetization using electric voltage only. The idea behind

large tilt of transmembrane helix 6 (TM-VI) and a smaller motion of helix 7 (TM-VII) at the inner surface of the cell membrane. These movements open a cleft that presents the ligand-binding site to the G protein. Scheerer and colleagues' results broadly match those of earlier biophysical studies[8], and agree quantitatively with those of work[9] that used pairs of 'spin labels' — which contain unpaired electrons that can be detected spectroscopically — to map out helix movement in activated rhodopsin. On the other hand, the authors' structure contrasts with a previously reported crystal structure[10] of activated rhodopsin (in the absence of a G protein) that differed little from the inactive state. Given the extensive biochemical and biophysical data in support of Scheerer and colleagues' structure for the activated state, it now seems likely that the previously reported structure[10] was modified by the conditions used to stabilize the crystals.

It might seem strange that the activating ligand, all-*trans* retinal, is not present in Scheerer and colleagues' opsin complex. But it is well established that active conformations of unliganded receptors, including opsin[11], exist in equilibrium with the inactive state. Indeed, 7TM receptors are often observed signalling with high activity through G proteins in the absence of agonist ligands. The crucial point in Scheerer and colleagues' structure[2] is that the observed active form is stabilized by a large excess of the G-protein fragment (the Gα peptide). It is the presence of this peptide that identifies the structure as a true active conformation. What is more surprising is that, in the absence of both all-*trans* retinal and the Gα peptide, opsin crystallizes in the same conformation as the authors report here (as reported earlier this year[12] by the same group).

The X-ray structure and orientation of the bound Gα peptide[2] are almost identical to those previously determined by nuclear magnetic resonance (NMR) studies[13] on solutions of a similar peptide interacting with activated rhodopsin. But Scheerer and colleagues' structure now shows how the peptide is stabilized by interactions with specific amino-acid residues that are exposed by the movement of TM-VI. Of particular note is that the peptide makes direct contact with the most evolutionarily conserved residue in 7TM receptors, an arginine residue at the intracellular end of TM-III.

The mechanism underlying the molecular recognition between opsin and the Gα peptide will probably be a good model for the interactions between 7TM receptors and the α-subunits of G proteins in general. The selectivity of these interactions is largely determined by the five carboxy-terminal residues of the α-subunits[14]. But it has been unexpectedly difficult to identify the corresponding selectivity-determining footprint on the receptors. The structure of Scheerer and colleagues' complex will therefore undoubtedly spur a wave of studies aimed at understanding this recognition process in detail. Moreover, interactions of 7TM receptors with G proteins, and their subsequent activation processes, are known to involve the recognition of additional structural elements on both the receptor and the G protein (Fig. 1). These processes can now be investigated in structure-based studies, as suggested by Scheerer and co-workers[2].

The observed conformation of activated opsin on the intracellular side of the cell membrane probably represents the activated state for 7TM receptors in general. But the relatively minor conformational changes that occur in the ligand-binding domain — which lies between the extracellular segments of the transmembrane helices — is likely to be unique to opsin. In rhodopsin, part of the extracellular region of opsin forms a protein 'plug' that fills the entrance to the main ligand-binding pocket, preventing movement of the transmembrane helices (Fig. 1). In contrast, in the previously determined structures of adrenergic receptors[5–7], the main ligand-binding crevice forms an open funnel towards the extracellular space. This allows ligands to diffuse in and out of the binding site, and could also allow inward movement of the extracellular segments of the transmembrane helices so that they surround the ligand. Such movement might aid ligand binding, and form part of a global 'toggle-switch' mechanism of receptor activation[8]. To test this possibility, we will need structures of 7TM receptors in complex not only with the G protein, but also with small-molecule agonists. ■

Thue W. Schwartz is at the Laboratory for Molecular Pharmacology, University of Copenhagen, Blegdamsvej 3, 2200 Copenhagen, Denmark. Wayne L. Hubbell is at the Jules Stein Eye Institute and the Department of Chemistry and Biochemistry, University of California, Los Angeles, Los Angeles, California 90095, USA.
e-mails: tws@sund.ku.dk; hubbellw@jsei.ucla.edu

1. Pierce, K. L., Premont, R. T. & Lefkowitz, R. J. *Nature Rev. Mol. Cell Biol.* **3,** 639–650 (2002).
2. Scheerer, P. *et al. Nature* **455,** 497–502 (2008).
3. Schertler, G. F. *Curr. Opin. Struct. Biol.* **15,** 408–415 (2005).
4. Rasmussen, S. G. F. *et al. Nature* **450,** 383–387 (2007).
5. Rosenbaum, D. M. *et al. Science* **318,** 1266–1273 (2007).
6. Cherezov, V. *et al. Science* **318,** 1258–1265 (2007).
7. Warne, T. *et al. Nature* **454,** 486–491 (2008).
8. Schwartz, T. W., Frimurer, T. M., Holst, B., Rosenkilde, M. M. & Elling, C. E. *Annu. Rev. Pharmacol. Toxicol.* **46,** 481–519 (2006).
9. Altenbach, C., Kusnetzow, A. K., Ernst, O. P., Hofmann, K. P. & Hubbell, W. L. *Proc. Natl Acad. Sci. USA* **105,** 7439–7444 (2008).
10. Salom, D. *et al. Proc. Natl Acad. Sci. USA* **103,** 16123–16128 (2006).
11. Woodruff, M. L. *et al. Nature Genet.* **35,** 158–164 (2003).
12. Park, J. H. *et al. Nature* **454,** 183–187 (2008).
13. Koenig, B. W. *et al. J. Mol. Biol.* **322,** 441–461 (2002).
14. Conklin, B. R., Farzel, Z., Lustig, K. D., Julius, D. & Bourne, H. R. *Nature* **363,** 274–276 (1993).

Competing financial interests: declared (see online article for details).

SOLID-STATE PHYSICS

New order for magnetism

Eiji Saitoh

Physicists have come up with an innovative way of manipulating the direction of magnetization in a solid. The approach might be used to make low-power-consumption computer memory devices.

On page 515 of this issue, Chiba *et al.*[1] show how the direction of magnetization in a solid can be controlled electrically. The significance of this achievement lies in the fact that magnetic data-storage and memory devices, such as hard disks, depend on manipulating the direction of magnetization. But methods for doing this have so far been inefficient.

The fundamental property of a magnet is that, in a magnetic field, it aligns in the direction of the field (Fig. 1a). A magnet consists of atomic-scale magnets, called spins, and if a strong magnetic field is applied to a fixed magnet, the spins rotate to align with the field. As a result, the magnetization in the magnet — the average of the spin-magnetic moment — takes on the direction of the applied magnetic field (Fig. 1b). Magnetic disks exploit this principle. In a hard disk, for instance, information is stored on a disk-shaped magnet in the form of patterns of local magnetization and, to write information, a pulse of current is applied to a small electromagnet that scans the disk. But this method is indirect: a magnetic field exists between the current and the manipulated magnetization, a situation that wastes energy.

A promising alternative would be to apply a spin-polarized current directly to the magnet instead of using current to generate the magnetic field (Fig. 1c). Such a current would exert torque on the magnetization by exchanging spin-angular momentum with it as it passes through the magnet. Exploitation of this phenomenon, called spin torque[2], is expected to allow the development of compact magnetic memory devices that can run on low power consumption. The use of multiferroics[3], materials in which ferroelectricity and ferromagnetism coexist, has also been proposed. But direct control of the magnetization direction has not yet been achieved in such systems.

Chiba *et al.*[1] demonstrate a new way of manipulating the direction of magnetization using electric voltage only. The idea behind

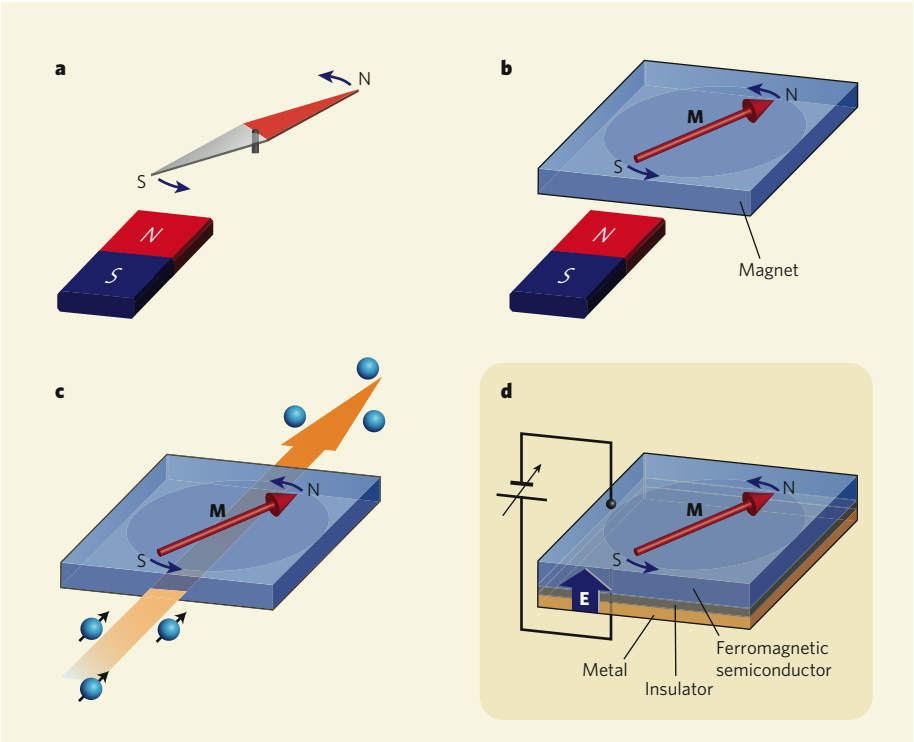


Figure 1 | Manipulation of magnetism. a, When a magnetic field is applied, a magnet rotates to align with the field. **b**, The magnetization (M) in the magnet can be rotated by applying an external magnetic field. Today's magnetic recording technology exploits this property. **c**, The direction of magnetization in the magnet is changed by applying a spin-polarized current (the spin-torque effect). **d**, Application of an electric field E rotates the direction of magnetization: Chiba *et al.*[1] demonstrate this effect experimentally with a ferromagnetic semiconductor, (GaMn)As.

their experiments is to control the magnetic anisotropy of a magnet and so generate preferential directions for the magnetization. This results from an interaction between the electron's spin and its motion (the spin–orbit interaction), which means that the electron's energy depends on the direction of magnetization. Such dependence is in general affected by the electronic structure — that is, the configuration and density of electrons in the magnet. The implication is that the magnetic anisotropy can be controlled by regulating the electronic structure.

To do this, Chiba *et al.* modulated the density of electrons using a metal–insulator–semiconductor device involving a semiconductor — a (GaMn)As film — that has ferromagnetic properties at low temperatures (Fig. 1d). The device includes a 'gate' electrode isolated electrically from the (GaMn)As film. When a negative voltage is applied to the gate electrode, carriers in the film that have positive charge (electron 'holes') are attracted towards the electrode. In contrast, when a positive voltage is applied, the number of electron holes near the gate electrode is reduced. This property allows the density of the electron holes, and thus the magnetic anisotropy in the (GaMn)As film beneath the electrode, to be controlled electrically, resulting in a change in the magnetization direction.

To determine the direction of magnetization, Chiba *et al.*[1] used the transport properties of carriers in (GaMn)As, measuring the voltage differences at the ends of the film when an electric current flows through it. These voltage differences reflect the relative angle between the current and the magnetization direction. The authors applied a weak external magnetic field to the (GaMn)As film and measured the voltage while changing the direction of the field.

They found that the magnetization direction in the film indeed deviates from the direction of the magnetic field. Moreover, the strength of this deviation varies with the carriers' density and the voltage applied to the gate electrode. The authors attribute this deviation to the magnetic anisotropy, and conclude that their result is evidence for voltage-induced modulation of the magnetic anisotropy. From the estimated anisotropy, and in the absence of an external magnetic field, the magnetization direction rotates 10° when the gate voltage changes from –12 volts to about 9 volts. In other words, electric manipulation of the magnetization direction is achieved.

The method proposed by Chiba *et al.*[1] does not require currents flowing in the device, but only the application of voltage. It is thus highly compatible with the existing metal-oxide semiconductor technology used in microprocessors and random-access memory devices. The principle involved here — that magnetic anisotropy depends on electronic structure — is also applicable to a wide range of materials, such as the surfaces of metallic or metal-oxide ultra-thin films[4]. Magnetic random-access memory devices made of metallic magnets are now being produced, and it is certainly worth exploring the possibility of electric control of magnetization also in metallic systems. ■

Eiji Saitoh is in the Department of Applied Physics and Physico-Informatics, Keio University, 3-14-1 Hiyoshi, Yokohama 223-8522, Japan.
e-mail: saitoheiji@appi.keio.ac.jp

1. Chiba, D. *et al. Nature* **455,** 515–518 (2008).
2. Slonczewski, J. C. *J. Magn. Magn. Mater.* **159,** L1–L7 (1996).
3. Tokura, Y. *Science* **312,** 1481–1482 (2006).
4. Weisheit, M. *et al. Science* **315,** 349–351 (2007).

HEARING

Route to authentic hair cells

Mats Ulfendahl

Existing therapies for hearing defects are generally ineffective in severe forms of deafness. A technical feat that generates sound-sensing hair cells in the inner ear of mice might have long-term potential.

Impaired hearing affects the communication abilities of approximately 17% of the US adult population[1]. Hearing defects of the outer or middle ear are generally treatable. By contrast, defects in the less accessible regions, such as the cochlea (the sound-sensing system of the inner ear) or the auditory nerve, are often permanent and have been considered untreatable. Such impairment, which is usually caused by excessive noise, ageing or genetic factors, involves degeneration of the mechanosensory receptors of the hearing organ — the inner and outer hair cells. On page 537 of this issue, Gubbels and colleagues[2] demonstrate that gene transfer to the inner-ear precursor cells of developing mouse embryos can be used to induce the formation of functional hair cells.

For decades, the only therapeutic option for hearing impairment was the use of hearing aids. These devices amplify and filter the incoming sound, improving the hearing threshold. But even with their use, hearing is far from normal, as the ear's receiving elements — the sensory hair cells — still do not function adequately. The introduction of cochlear implants in the 1980s was a big step forward, as they completely bypass the defective sensory cells and electrically stimulate the next

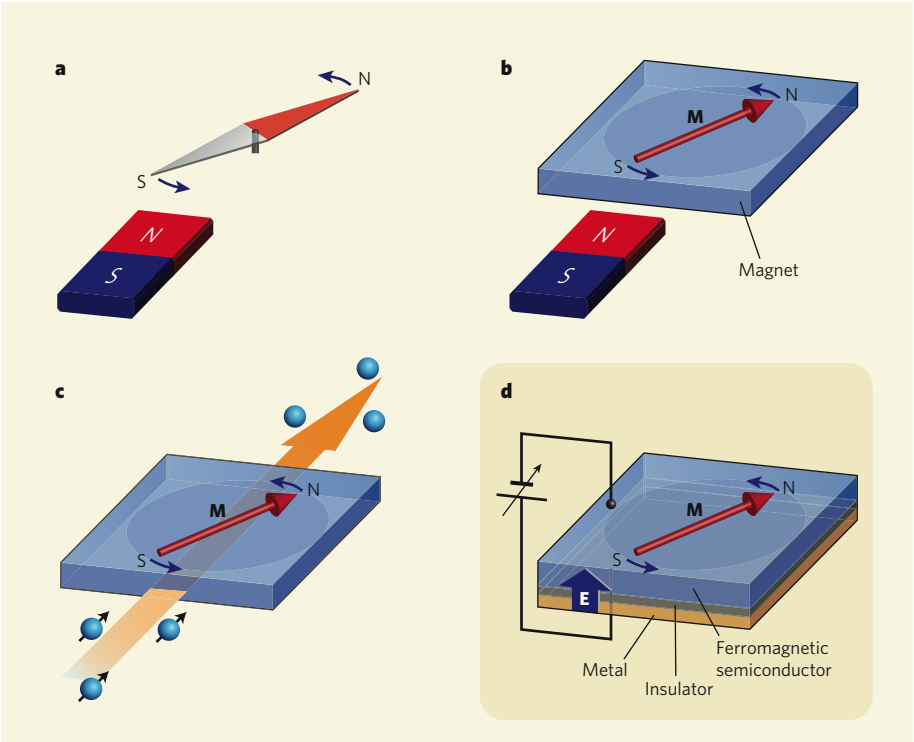


Figure 1 | Manipulation of magnetism. a, When a magnetic field is applied, a magnet rotates to align with the field. **b**, The magnetization (M) in the magnet can be rotated by applying an external magnetic field. Today's magnetic recording technology exploits this property. **c**, The direction of magnetization in the magnet is changed by applying a spin-polarized current (the spin-torque effect). **d**, Application of an electric field E rotates the direction of magnetization: Chiba *et al.*[1] demonstrate this effect experimentally with a ferromagnetic semiconductor, (GaMn)As.

their experiments is to control the magnetic anisotropy of a magnet and so generate preferential directions for the magnetization. This results from an interaction between the electron's spin and its motion (the spin–orbit interaction), which means that the electron's energy depends on the direction of magnetization. Such dependence is in general affected by the electronic structure — that is, the configuration and density of electrons in the magnet. The implication is that the magnetic anisotropy can be controlled by regulating the electronic structure.

To do this, Chiba *et al.* modulated the density of electrons using a metal–insulator–semiconductor device involving a semiconductor — a (GaMn)As film — that has ferromagnetic properties at low temperatures (Fig. 1d). The device includes a 'gate' electrode isolated electrically from the (GaMn)As film. When a negative voltage is applied to the gate electrode, carriers in the film that have positive charge (electron 'holes') are attracted towards the electrode. In contrast, when a positive voltage is applied, the number of electron holes near the gate electrode is reduced. This property allows the density of the electron holes, and thus the magnetic anisotropy in the (GaMn)As film beneath the electrode, to be controlled electrically, resulting in a change in the magnetization direction.

To determine the direction of magnetization, Chiba *et al.*[1] used the transport properties of carriers in (GaMn)As, measuring the voltage differences at the ends of the film when an electric current flows through it. These voltage differences reflect the relative angle between the current and the magnetization direction. The authors applied a weak external magnetic field to the (GaMn)As film and measured the voltage while changing the direction of the field.

They found that the magnetization direction in the film indeed deviates from the direction of the magnetic field. Moreover, the strength of this deviation varies with the carriers' density and the voltage applied to the gate electrode. The authors attribute this deviation to the magnetic anisotropy, and conclude that their result is evidence for voltage-induced modulation of the magnetic anisotropy. From the estimated anisotropy, and in the absence of an external magnetic field, the magnetization direction rotates 10° when the gate voltage changes from –12 volts to about 9 volts. In other words, electric manipulation of the magnetization direction is achieved.

The method proposed by Chiba *et al.*[1] does not require currents flowing in the device, but only the application of voltage. It is thus highly compatible with the existing metal-oxide semiconductor technology used in microprocessors and random-access memory devices. The principle involved here — that magnetic anisotropy depends on electronic structure — is also applicable to a wide range of materials, such as the surfaces of metallic or metal-oxide ultra-thin films[4]. Magnetic random-access memory devices made of metallic magnets are now being produced, and it is certainly worth exploring the possibility of electric control of magnetization also in metallic systems. ■

Eiji Saitoh is in the Department of Applied Physics and Physico-Informatics, Keio University, 3-14-1 Hiyoshi, Yokohama 223-8522, Japan.
e-mail: saitoheiji@appi.keio.ac.jp

1. Chiba, D. *et al. Nature* **455,** 515–518 (2008).
2. Slonczewski, J. C. *J. Magn. Magn. Mater.* **159,** L1–L7 (1996).
3. Tokura, Y. *Science* **312,** 1481–1482 (2006).
4. Weisheit, M. *et al. Science* **315,** 349–351 (2007).

HEARING

Route to authentic hair cells

Mats Ulfendahl

Existing therapies for hearing defects are generally ineffective in severe forms of deafness. A technical feat that generates sound-sensing hair cells in the inner ear of mice might have long-term potential.

Impaired hearing affects the communication abilities of approximately 17% of the US adult population[1]. Hearing defects of the outer or middle ear are generally treatable. By contrast, defects in the less accessible regions, such as the cochlea (the sound-sensing system of the inner ear) or the auditory nerve, are often permanent and have been considered untreatable. Such impairment, which is usually caused by excessive noise, ageing or genetic factors, involves degeneration of the mechanosensory receptors of the hearing organ — the inner and outer hair cells. On page 537 of this issue, Gubbels and colleagues[2] demonstrate that gene transfer to the inner-ear precursor cells of developing mouse embryos can be used to induce the formation of functional hair cells.

For decades, the only therapeutic option for hearing impairment was the use of hearing aids. These devices amplify and filter the incoming sound, improving the hearing threshold. But even with their use, hearing is far from normal, as the ear's receiving elements — the sensory hair cells — still do not function adequately. The introduction of cochlear implants in the 1980s was a big step forward, as they completely bypass the defective sensory cells and electrically stimulate the next

functional level in the auditory pathway, the auditory neurons. Moreover, the successful use of implants indicated that more invasive treatment of the inner ear, for example at the cellular level, might be feasible.

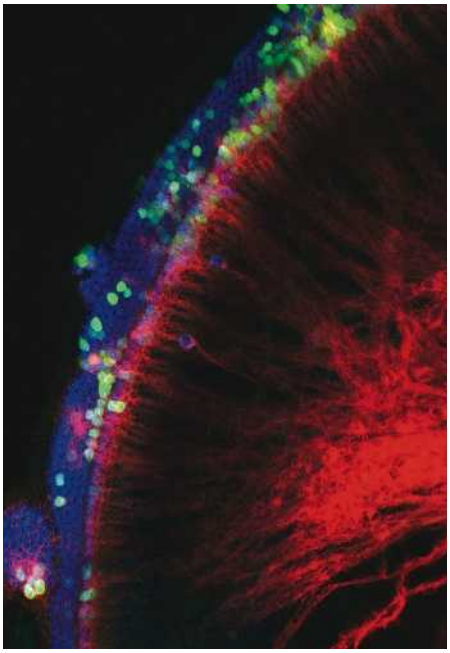
Four main strategies are currently being explored for dealing with defects of the inner ear: prevention, intervention, substitution and regeneration. The most straightforward of these is prevention, in which protective treatment is started before an anticipated trauma. But such treatment is essentially limited to situations in which, for example, a drug with known auditory side effects, such as the anticancer drug cisplatin, is to be used. Intervention therapy after trauma involves modifying the biological processes involved in the damage; but this is feasible only for acute trauma such as that following accidental exposure to intense and damaging noise .

To treat permanent damage to the inner ear, completely different strategies must be explored. Substitution therapy, which aims to replace missing or dysfunctional cells or cellular components, is an exciting option, as is regeneration therapy, which involves inducing the differentiation of progenitor cells in the adult inner ear.

Previous substitution experiments, both *in vitro*[3] and in the intact inner ear[4,5], have shown that the formation of new hair cells from non-sensory cells in the inner ear can be triggered by increasing the expression levels of Atoh1, a gene transcription factor that is required for hair-cell development[6]. But *de novo* generation of hair cells does not mean that the cells are fully functional.

Gubbels *et al.*[2] have designed an elegant series of substitution/regeneration experiments, using the technique of transfection, to introduce the gene encoding Atoh1 *(Atoh1)* directly into inner-ear progenitor cells of mouse embryos at day 11.5 of development while still in the uterus. The *in utero* approach takes gene transfer to the inner ear a step forward, as it points to the possibility of treating hereditary defects at a very early stage, before degenerative changes become permanent. To follow the fate of successfully transfected cells, the gene encoding the green fluorescent protein marker was introduced with *Atoh1* (Fig. 1).

The transfected cells displayed their characteristic bundles of hairs. They also seemed to form the correct neural contacts with afferent nerve fibres, thus allowing the newly formed cells to interact with the auditory nervous system (Fig. 1). By mechanically stimulating the hair bundle while recording using electrophysiological methods, Gubbels *et al.* showed that the cells transfected with *Atoh1* were capable of mechano-electrical transduction — a fundamental process in auditory function that involves the conversion of biomechanical stimuli to electrical impulses in the auditory nerve fibres. These tests demonstrate that the cells are indeed functional auditory sensory cells.



S. P. GUBBELS, D. W. WOESSNER, C. BRESEE & J. V. BRIGANDE

Figure 1 | Engineered hair cells fit in. Gubbels *et al.*[2] introduced the gene encoding Atoh1 into precursor hair cells in the inner ear of mice at day 11.5 of embryonic development *in utero*. As shown here, by embryonic day 18.5 the transfected cells, which were labelled with a green fluorescent marker protein (green), had differentiated into hair cells and could be seen among untransfected, natural hair cells. Moreover, like natural hair cells, they formed appropriate connections with nerve fibres (red).

Gubbels and colleagues' findings[2], together with earlier work[3–5] on gene transfer in the inner ear, provide proof-of-principle that the main elements in the inner ear can be formed *de novo*. These observations, and the *in utero* technique used by Gubbels *et al.* for gene transfer, will have far-reaching consequences, opening the way to the development of totally new therapies for hearing loss due to defects of the sensory hair cells.

The identification of genetic defects underlying many forms of hereditary deafness has stimulated growing interest in therapeutic methods based on gene transfer. For example, in some parts of the world, mutations in the gene encoding connexin 26 — an essential component of gap junctions between cells — account for more than 50% of cases of severe hereditary deafness not associated with other symptoms[7]. Whether the aim is to restore normal connexin-26 levels or to ensure the correct formation of sensory cells during development, exploring methods for *in vivo* gene transfer should be useful.

But as exciting as these results are, several issues surrounding a gene-transfer approach must first be considered. For example, if the technique were to be applied prenatally in humans, the risks for both the fetus and the mother are clearly a limiting factor. Moreover, inner-ear gene-transfer protocols reported so far have been effective for only a limited period of time. Nevertheless, although it will be some years before Gubbels and colleagues' techniques are translated into clinical practice, the future certainly looks promising. ■

Mats Ulfendahl is at the Center for Hearing and Communication Research, and the Department of Clinical Neuroscience, Karolinska Institutet, Building M1, Karolinska University Hospital, SE-171 76 Stockholm, Sweden.
e-mail: mats.ulfendahl@ki.se

1. www.nidcd.nih.gov/health/statistics/quick.htm
2. Gubbels, S. P., Woessner, D. W., Mitchell, J. C., Ricci, A. J. & Brigande, J. V. *Nature* **455,** 537–541 (2008).
3. Zheng, J. L. & Gao, W. Q. *Nature Neurosci.* **3,** 580–586 (2000).
4. Kawamoto, K. *et al. J. Neurosci.* **23,** 4395–4400 (2003).
5. Izumikawa, M. *et al. Nature Med.* **11,** 271–276 (2005).
6. Bermingham, N. A. *et al. Science* **284,** 1837–1841 (1999).
7. Sabag, A. D., Dagan, O. & Avraham, K. B. *J. Basic Clin. Physiol. Pharmacol.* **16,** 101–116 (2005).

ASTROPHYSICS

How fast can you blink?

Chryssa Kouveliotou

Serendipitous observations have revealed fast optical flaring after the onset of X-ray-burst activity from a source in our Milky Way galaxy. It could be the first time this has been observed in a rare kind of neutron star.

Magnetars are slowly rotating neutron stars with extremely strong magnetic fields. Originally classified as members of two distinct families of astronomical objects, soft γ-ray repeaters (SGRs) and anomalous X-ray pulsars (AXPs), they were merged into a single class when it was recognized that they share several properties. Some members of other groups of celestial oddities, including dim isolated neutron stars (DINs), have recently also joined their ranks. But none of them has exhibited the fast optical flaring activity of the new candidate magnetar, SWIFT J195509.6+261406, described by Castro-Tirado *et al.*[1] and Stefanescu *et al.*[2] in this issue.

The main question to arise from these observations is whether this discovery is the result of serendipitous observational circumstances or part and parcel of magnetar properties in general. Several factors make this a difficult, if

functional level in the auditory pathway, the auditory neurons. Moreover, the successful use of implants indicated that more invasive treatment of the inner ear, for example at the cellular level, might be feasible.

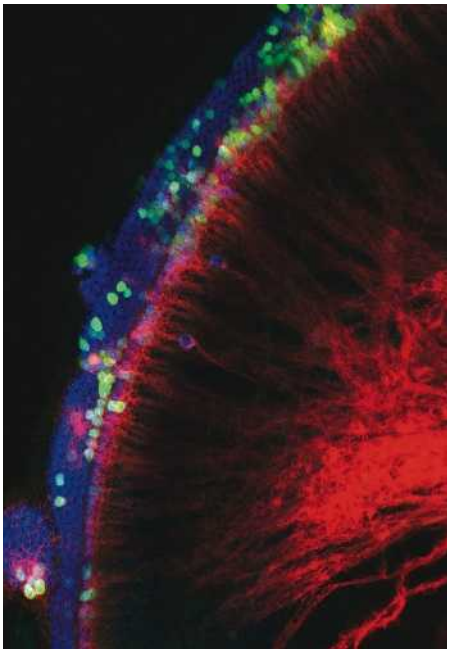
Four main strategies are currently being explored for dealing with defects of the inner ear: prevention, intervention, substitution and regeneration. The most straightforward of these is prevention, in which protective treatment is started before an anticipated trauma. But such treatment is essentially limited to situations in which, for example, a drug with known auditory side effects, such as the anticancer drug cisplatin, is to be used. Intervention therapy after trauma involves modifying the biological processes involved in the damage; but this is feasible only for acute trauma such as that following accidental exposure to intense and damaging noise .

To treat permanent damage to the inner ear, completely different strategies must be explored. Substitution therapy, which aims to replace missing or dysfunctional cells or cellular components, is an exciting option, as is regeneration therapy, which involves inducing the differentiation of progenitor cells in the adult inner ear.

Previous substitution experiments, both *in vitro*[3] and in the intact inner ear[4,5], have shown that the formation of new hair cells from non-sensory cells in the inner ear can be triggered by increasing the expression levels of Atoh1, a gene transcription factor that is required for hair-cell development[6]. But *de novo* generation of hair cells does not mean that the cells are fully functional.

Gubbels *et al.*[2] have designed an elegant series of substitution/regeneration experiments, using the technique of transfection, to introduce the gene encoding Atoh1 *(Atoh1)* directly into inner-ear progenitor cells of mouse embryos at day 11.5 of development while still in the uterus. The *in utero* approach takes gene transfer to the inner ear a step forward, as it points to the possibility of treating hereditary defects at a very early stage, before degenerative changes become permanent. To follow the fate of successfully transfected cells, the gene encoding the green fluorescent protein marker was introduced with *Atoh1* (Fig. 1).

The transfected cells displayed their characteristic bundles of hairs. They also seemed to form the correct neural contacts with afferent nerve fibres, thus allowing the newly formed cells to interact with the auditory nervous system (Fig. 1). By mechanically stimulating the hair bundle while recording using electrophysiological methods, Gubbels *et al.* showed that the cells transfected with *Atoh1* were capable of mechano-electrical transduction — a fundamental process in auditory function that involves the conversion of biomechanical stimuli to electrical impulses in the auditory nerve fibres. These tests demonstrate that the cells are indeed functional auditory sensory cells.



S. P. GUBBELS, D. W. WOESSNER, C. BRESEE & J. V. BRIGANDE

Figure 1 | Engineered hair cells fit in. Gubbels *et al.*[2] introduced the gene encoding Atoh1 into precursor hair cells in the inner ear of mice at day 11.5 of embryonic development *in utero*. As shown here, by embryonic day 18.5 the transfected cells, which were labelled with a green fluorescent marker protein (green), had differentiated into hair cells and could be seen among untransfected, natural hair cells. Moreover, like natural hair cells, they formed appropriate connections with nerve fibres (red).

Gubbels and colleagues' findings[2], together with earlier work[3–5] on gene transfer in the inner ear, provide proof-of-principle that the main elements in the inner ear can be formed *de novo*. These observations, and the *in utero* technique used by Gubbels *et al.* for gene transfer, will have far-reaching consequences, opening the way to the development of totally new therapies for hearing loss due to defects of the sensory hair cells.

The identification of genetic defects underlying many forms of hereditary deafness has stimulated growing interest in therapeutic methods based on gene transfer. For example, in some parts of the world, mutations in the gene encoding connexin 26 — an essential component of gap junctions between cells — account for more than 50% of cases of severe hereditary deafness not associated with other symptoms[7]. Whether the aim is to restore normal connexin-26 levels or to ensure the correct formation of sensory cells during development, exploring methods for *in vivo* gene transfer should be useful.

But as exciting as these results are, several issues surrounding a gene-transfer approach must first be considered. For example, if the technique were to be applied prenatally in humans, the risks for both the fetus and the mother are clearly a limiting factor. Moreover, inner-ear gene-transfer protocols reported so far have been effective for only a limited period of time. Nevertheless, although it will be some years before Gubbels and colleagues' techniques are translated into clinical practice, the future certainly looks promising. ■

Mats Ulfendahl is at the Center for Hearing and Communication Research, and the Department of Clinical Neuroscience, Karolinska Institutet, Building M1, Karolinska University Hospital, SE-171 76 Stockholm, Sweden.
e-mail: mats.ulfendahl@ki.se

1. www.nidcd.nih.gov/health/statistics/quick.htm
2. Gubbels, S. P., Woessner, D. W., Mitchell, J. C., Ricci, A. J. & Brigande, J. V. *Nature* **455,** 537–541 (2008).
3. Zheng, J. L. & Gao, W. Q. *Nature Neurosci.* **3,** 580–586 (2000).
4. Kawamoto, K. *et al. J. Neurosci.* **23,** 4395–4400 (2003).
5. Izumikawa, M. *et al. Nature Med.* **11,** 271–276 (2005).
6. Bermingham, N. A. *et al. Science* **284,** 1837–1841 (1999).
7. Sabag, A. D., Dagan, O. & Avraham, K. B. *J. Basic Clin. Physiol. Pharmacol.* **16,** 101–116 (2005).

ASTROPHYSICS

How fast can you blink?

Chryssa Kouveliotou

Serendipitous observations have revealed fast optical flaring after the onset of X-ray-burst activity from a source in our Milky Way galaxy. It could be the first time this has been observed in a rare kind of neutron star.

Magnetars are slowly rotating neutron stars with extremely strong magnetic fields. Originally classified as members of two distinct families of astronomical objects, soft γ-ray repeaters (SGRs) and anomalous X-ray pulsars (AXPs), they were merged into a single class when it was recognized that they share several properties. Some members of other groups of celestial oddities, including dim isolated neutron stars (DINs), have recently also joined their ranks. But none of them has exhibited the fast optical flaring activity of the new candidate magnetar, SWIFT J195509.6+261406, described by Castro-Tirado *et al.*[1] and Stefanescu *et al.*[2] in this issue.

The main question to arise from these observations is whether this discovery is the result of serendipitous observational circumstances or part and parcel of magnetar properties in general. Several factors make this a difficult, if

not an impossible, question to answer.

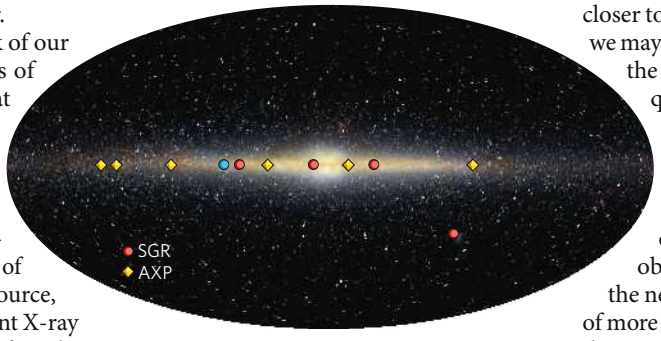
Most magnetars reside in the disk of our Galaxy (Fig. 1). Of these, members of the SGR subclass are at distances of at least 33,000 light years from Earth, and thus are subject to large optical extinction from gas and dust in the line of sight. In these cases, observations in the infrared part of the electromagnetic spectrum are the only means of detection. But thus far, just one SGR source, which was discovered following a giant X-ray outburst[3,4], has been detected in the infrared.

In contrast, AXPs are closer to Earth, 10,000–16,000 light years away, allowing both optical and infrared observations. But only one source, 4U 0142+61, has been detected in optical emission[5], which was found to pulse at the same period of 8.7 seconds detected in the persistent X-ray flux of this source. The discovery of the spin period clearly associated this emission with the AXP and identified it as magnetospheric in origin[6]. But the optical flux from 4U 0142+61 exhibited only coherent modulation and did not include any fast-flaring emission. In total, five AXPs (including 4U 0142+61) have been detected by their infrared emissions, showing a variability that does not always correlate with that in their X-ray band[7,8]. In fact, an anticorrelation or a lack of correlation between the two emissions has also been observed[9,10].

Given this motley crew of magnetar properties, it is not surprising that Stefanescu *et al.*[2] (page 503) find no correlation between the fast optical flaring and the X-ray flaring of SWIFT J195509.6+261406. The difficulty in identifying the nature of this source is enhanced by the fact that observers and theorists alike are in uncharted waters in many respects.

The frequency of the optical flares and their temporal characteristics were resolved for the first time owing to the unique temporal resolution (4 μs) of the OPTIMA camera on a 1.3-metre telescope at the Skinakas Observatory in Crete. The shortest timescales observed were 0.3–0.4 seconds — that's pretty fast blinking. Moreover, Stefanescu and colleagues were able to observe the source just 421 seconds after its first X-ray burst. The OPTIMA camera was following a previous signal from a different source that happened to be located close to SWIFT J195509.6+261406. In contrast, all other optical or infrared follow-up observations of magnetars were performed days after their activation in the X-ray band. Furthermore, the new candidate magnetar happened to be nearby, with an estimated distance of no more than 16,000 light years, significantly reducing light extinction in the line of sight. This value was confirmed by the observations of Castro-Tirado *et al.*[1] (page 506).

Taking into account the energetics of the X-ray and optical flares, and having eliminated most other options, it seems that SWIFT J195509.6+261406 is indeed a nearby magnetar. Thus, serendipitous circumstances paired with instrumental capabilities have opened a new window on the antics of these rare neutron stars.



E. L. WRIGHT (UCLA)/COBE PROJECT/DIRBE/NASA

Figure 1 | Candidate magnetars in the Milky Way. The image shows the distribution of two subclasses of ultra-magnetic neutron stars (magnetars) — soft γ-ray repeaters (SGRs) and anomalous X-ray pulsars (AXPs) — in the plane of our Galaxy. The red dot out of the plane indicates a source in the Large Magellanic Cloud, a nearby (about 180,000 light years away) satellite galaxy of the Milky Way. The new candidate magnetar, SWIFT J195509.6+261406, described by Castro-Tirado *et al.*[1] and Stefanescu *et al.*[2], is indicated by the blue dot.

Interestingly, and as Castro-Tirado *et al.*[1] point out, the quiescent X-ray luminosity of the new source is far below (by a factor of almost 100) that of all known SGRs and AXPs, and is closer to the luminosities of DINs. Therefore, we may be observing a magnetar on its way to the neutron-star graveyard. This begs the question: what is the physical origin of such fast optical emission in magnetars? Does it occur in SGRs and AXPs, or is it a property of DINs alone? Alternatively, is it an attribute of just one, odd celestial object? In an era of observatories such as Swift and Fermi, the net is being cast wide for the discovery of more magnetars. One would also hope that the OPTIMA camera will still be up and running and looking at the right part of the sky again. ■

Chryssa Kouveliotou is at the NASA Marshall Space Flight Center, Huntsville, Alabama 35805, USA.
e-mail: chryssa.kouveliotou@nasa.gov

1. Castro-Tirado, A. J. *et al. Nature* **455,** 506–509 (2008).
2. Stefanescu, A. *et al. Nature* **455,** 503–505 (2008).
3. Israel, G. L. *et al. Astrophys. J.* **438,** L1–L4 (2005).
4. Kosugi, G., Ogasawara, R. & Terada, H. *Astrophys. J.* **623,** L125–L128 (2005).
5. Hulleman, F. *et al. Astrophys. J.* **563,** L49–L52 (2001).
6. Kern, B. & Martin, C. *Nature* **417,** 527–529 (2002).
7. Tam, C. R., Kaspi, V. M., van Kerkwijk, M. H. & Durant, M. *Astrophys. J.* **617,** L53–L56 (2004).
8. Gotthelf, E., Halpern, J., Buxton, M. & Bailyn, C. *Astrophys. J.* **605,** 368–377 (2004).
9. Durant, M. & van Kerkwijk, M. *Astrophys. J.* **627,** 376–382 (2005).
10. Testa, V. *et al. Astron. Astrophys.* **482,** 607–615 (2008).

NEUROSCIENCE

An ageing view of myelin repair

Klaus-Armin Nave

When the myelin layer that covers neuronal processes is lost through disease, neural stem cells recapitulate the developmental program of 'myelination'. The underlying molecular mechanisms often fail in the ageing brain.

In the brain, the function of neurons and their long axonal processes depends on neighbouring support cells called glia. For example, one class of glia, called oligodendrocytes, wrap the neurons' axons in a multilayered spiral extension of their own cell membrane — the myelin sheath. This insulating layer is required not only for fast nerve-impulse conduction along a neuron, and so for virtually all motor, sensory and higher functions of the brain, but also for the integrity and long-term survival of the axons. Myelin can be damaged in various diseases, including multiple sclerosis, but the mammalian brain has a remarkable ability to repair itself. With age, however, the efficiency of this remyelination declines or is lost completely. Writing in *Nature Neuroscience*, Shen *et al.*[1] provide insights into how this failure of remyelination can be explained at a molecular level.

In multiple sclerosis, autoreactive immune cells repeatedly attack oligodendrocytes and myelin, creating local areas of inflammation, called plaques, in the brain's white matter. Normally, resident oligodendrocyte precursor (OP) cells repopulate the affected area and differentiate into mature oligodendrocytes, which remyelinate the damaged axons. But if this process is inefficient, the affected axons — those that have survived the acute phase of inflammation — tend to degenerate[2].

Studies on demyelination have been carried out both in mice with experimental allergic encephalomyelitis, which resembles the inflammation of multiple sclerosis, and in animals with chemically induced demyelination[3]. The results indicate that, in older animals, myelin repair is far less efficient than in younger mice with the same degree of demyelination. Why does the brain's ability to remyelinate lesions decline with age? Shen *et al.*[1] find that, at least in part, altered 'epigenetic' regulation of gene expression might be the reason.

Epigenetic mechanisms reflect changes in gene expression that are not dictated by the

not an impossible, question to answer.

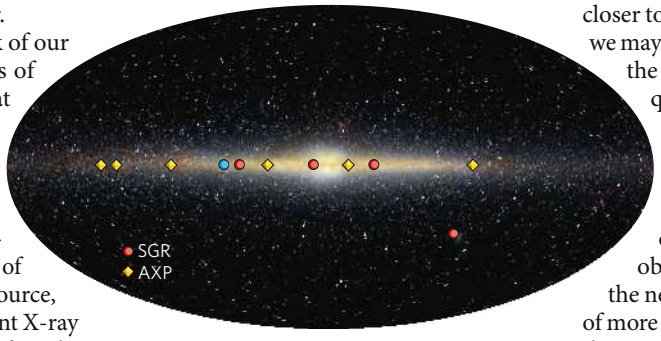
Most magnetars reside in the disk of our Galaxy (Fig. 1). Of these, members of the SGR subclass are at distances of at least 33,000 light years from Earth, and thus are subject to large optical extinction from gas and dust in the line of sight. In these cases, observations in the infrared part of the electromagnetic spectrum are the only means of detection. But thus far, just one SGR source, which was discovered following a giant X-ray outburst[3,4], has been detected in the infrared.

In contrast, AXPs are closer to Earth, 10,000–16,000 light years away, allowing both optical and infrared observations. But only one source, 4U 0142+61, has been detected in optical emission[5], which was found to pulse at the same period of 8.7 seconds detected in the persistent X-ray flux of this source. The discovery of the spin period clearly associated this emission with the AXP and identified it as magnetospheric in origin[6]. But the optical flux from 4U 0142+61 exhibited only coherent modulation and did not include any fast-flaring emission. In total, five AXPs (including 4U 0142+61) have been detected by their infrared emissions, showing a variability that does not always correlate with that in their X-ray band[7,8]. In fact, an anticorrelation or a lack of correlation between the two emissions has also been observed[9,10].

Given this motley crew of magnetar properties, it is not surprising that Stefanescu *et al.*[2] (page 503) find no correlation between the fast optical flaring and the X-ray flaring of SWIFT J195509.6+261406. The difficulty in identifying the nature of this source is enhanced by the fact that observers and theorists alike are in uncharted waters in many respects.

The frequency of the optical flares and their temporal characteristics were resolved for the first time owing to the unique temporal resolution (4 μs) of the OPTIMA camera on a 1.3-metre telescope at the Skinakas Observatory in Crete. The shortest timescales observed were 0.3–0.4 seconds — that's pretty fast blinking. Moreover, Stefanescu and colleagues were able to observe the source just 421 seconds after its first X-ray burst. The OPTIMA camera was following a previous signal from a different source that happened to be located close to SWIFT J195509.6+261406. In contrast, all other optical or infrared follow-up observations of magnetars were performed days after their activation in the X-ray band. Furthermore, the new candidate magnetar happened to be nearby, with an estimated distance of no more than 16,000 light years, significantly reducing light extinction in the line of sight. This value was confirmed by the observations of Castro-Tirado *et al.*[1] (page 506).

Taking into account the energetics of the X-ray and optical flares, and having eliminated most other options, it seems that SWIFT J195509.6+261406 is indeed a nearby magnetar. Thus, serendipitous circumstances paired with instrumental capabilities have opened a new window on the antics of these rare neutron stars.



E. L. WRIGHT (UCLA)/COBE PROJECT/DIRBE/NASA

Figure 1 | Candidate magnetars in the Milky Way. The image shows the distribution of two subclasses of ultra-magnetic neutron stars (magnetars) — soft γ-ray repeaters (SGRs) and anomalous X-ray pulsars (AXPs) — in the plane of our Galaxy. The red dot out of the plane indicates a source in the Large Magellanic Cloud, a nearby (about 180,000 light years away) satellite galaxy of the Milky Way. The new candidate magnetar, SWIFT J195509.6+261406, described by Castro-Tirado *et al.*[1] and Stefanescu *et al.*[2], is indicated by the blue dot.

Interestingly, and as Castro-Tirado *et al.*[1] point out, the quiescent X-ray luminosity of the new source is far below (by a factor of almost 100) that of all known SGRs and AXPs, and is closer to the luminosities of DINs. Therefore, we may be observing a magnetar on its way to the neutron-star graveyard. This begs the question: what is the physical origin of such fast optical emission in magnetars? Does it occur in SGRs and AXPs, or is it a property of DINs alone? Alternatively, is it an attribute of just one, odd celestial object? In an era of observatories such as Swift and Fermi, the net is being cast wide for the discovery of more magnetars. One would also hope that the OPTIMA camera will still be up and running and looking at the right part of the sky again. ■

Chryssa Kouveliotou is at the NASA Marshall Space Flight Center, Huntsville, Alabama 35805, USA.
e-mail: chryssa.kouveliotou@nasa.gov

1. Castro-Tirado, A. J. *et al. Nature* **455,** 506–509 (2008).
2. Stefanescu, A. *et al. Nature* **455,** 503–505 (2008).
3. Israel, G. L. *et al. Astrophys. J.* **438,** L1–L4 (2005).
4. Kosugi, G., Ogasawara, R. & Terada, H. *Astrophys. J.* **623,** L125–L128 (2005).
5. Hulleman, F. *et al. Astrophys. J.* **563,** L49–L52 (2001).
6. Kern, B. & Martin, C. *Nature* **417,** 527–529 (2002).
7. Tam, C. R., Kaspi, V. M., van Kerkwijk, M. H. & Durant, M. *Astrophys. J.* **617,** L53–L56 (2004).
8. Gotthelf, E., Halpern, J., Buxton, M. & Bailyn, C. *Astrophys. J.* **605,** 368–377 (2004).
9. Durant, M. & van Kerkwijk, M. *Astrophys. J.* **627,** 376–382 (2005).
10. Testa, V. *et al. Astron. Astrophys.* **482,** 607–615 (2008).

NEUROSCIENCE

An ageing view of myelin repair

Klaus-Armin Nave

When the myelin layer that covers neuronal processes is lost through disease, neural stem cells recapitulate the developmental program of 'myelination'. The underlying molecular mechanisms often fail in the ageing brain.

In the brain, the function of neurons and their long axonal processes depends on neighbouring support cells called glia. For example, one class of glia, called oligodendrocytes, wrap the neurons' axons in a multilayered spiral extension of their own cell membrane — the myelin sheath. This insulating layer is required not only for fast nerve-impulse conduction along a neuron, and so for virtually all motor, sensory and higher functions of the brain, but also for the integrity and long-term survival of the axons. Myelin can be damaged in various diseases, including multiple sclerosis, but the mammalian brain has a remarkable ability to repair itself. With age, however, the efficiency of this remyelination declines or is lost completely. Writing in *Nature Neuroscience*, Shen *et al.*[1] provide insights into how this failure of remyelination can be explained at a molecular level.

In multiple sclerosis, autoreactive immune cells repeatedly attack oligodendrocytes and myelin, creating local areas of inflammation, called plaques, in the brain's white matter. Normally, resident oligodendrocyte precursor (OP) cells repopulate the affected area and differentiate into mature oligodendrocytes, which remyelinate the damaged axons. But if this process is inefficient, the affected axons — those that have survived the acute phase of inflammation — tend to degenerate[2].

Studies on demyelination have been carried out both in mice with experimental allergic encephalomyelitis, which resembles the inflammation of multiple sclerosis, and in animals with chemically induced demyelination[3]. The results indicate that, in older animals, myelin repair is far less efficient than in younger mice with the same degree of demyelination. Why does the brain's ability to remyelinate lesions decline with age? Shen *et al.*[1] find that, at least in part, altered 'epigenetic' regulation of gene expression might be the reason.

Epigenetic mechanisms reflect changes in gene expression that are not dictated by the

DNA sequence itself. They include changes in chromatin — the complexes of DNA with histone proteins. For example, chromatin remodelling involves histone deacetylase (HDAC) enzymes, which remove acetyl groups from the lysine amino-acid residues in histone proteins. The deacetylated histones favour a more condensed chromatin conformation, which is responsible for inactivation of transcription. Also, the deacetylated histones themselves may have a reduced affinity for some transcription factors.

Shen and colleagues' previous work[4] established that, during normal development, myelination requires a reduced activity of gene transcription factors that inhibit oligodendrocyte differentiation and of those that cause OP cells to retain their stem-cell character. They also found that HDACs are essential for the differentiation of OP cells, because inhibiting these enzymes with valproic acid dramatically interfered with normal development. The authors then looked for a link between the apparent decline in remyelination with age and changes in the gene expression patterns of OP cells in adult animals.

To do this, they now examine[1] a well-established mouse model of demyelination that is generated by giving animals the toxin cuprizone. In cuprizone-treated young adult mice (eight weeks old), demyelination is followed by spontaneous remyelination. The authors find that remyelination was associated with a complex pattern of changes in gene expression that was similar to that observed during myelination in newborns. When the mice were given the HDAC inhibitor valproic acid while on a cuprizone diet, the differentiation of OP cells, and thus remyelination of white matter, was markedly impaired. Valproic acid had no obvious effect on other cell types, or on the severity of inflammation or on axon survival.

Cuprizone induced a similar pattern of demyelination in older (10-month-old) mice, but the subsequent remyelination was not accompanied by the 'molecular signature' detected in young mice. Therefore, impaired remyelination was associated with an age-dependent loss of 'epigenetic memory'. This was strikingly similar to the results seen in younger mice treated with both valproic acid and cuprizone, in particular with respect to a reduction in the expression of HDAC1 and HDAC8 by OP cells.

A closer look revealed that the impaired remyelination in older mice did not result from various possible effects — an altered inflammatory response, for example, or an abnormal increase in the number of 'reactive' astrocytes (a known side effect of brain injury), or axonal degeneration at the lesion site. Rather, genes encoding inhibitory transcription factors (such as Hes5) and factors that promote stem-cell activity (such as Sox2) were transcribed for much longer than normally seen in remyelination. Moreover, the enzyme RNA Pol II, which mediates transcription, was physically associated with these genetic loci. Thus, the ability to recruit HDACs to the promoter regions of inhibitory transcription factors, which is seen in young mice, seems to be lost in older animals.

To identify the HDACs that shut down the various inhibitors of differentiation, Shen and colleagues[1] performed gene-silencing experiments in oligodendrocytes grown in culture. The exact molecular code revealed will be of interest only to the specialist, but the message will be of great interest to stem-cell biologists and researchers in the fields of ageing and myelin-related disease: age-dependent differences in the brain's ability to repair a lesion are due to variations in the epigenetic regulation of gene expression in OP cells.

The proposed link between failing HDAC activity in the OP cells of older mice and inefficient remyelination hint at the need for 'gain-of-function' experiments. If altered HDAC activity is indeed responsible for poor remyelination, transgenic overexpression of HDAC1 in OP cells should overcome this.

But it would be too simplistic to assume that loss of epigenetic memory is the only mechanism behind poor remyelination. Various proteins also seem to directly inhibit remyelination by oligodendrocytes. For example, the cell-adhesion protein PSA-NCAM is expressed in demyelinated axons[5]. Also, LINGO-1, which was originally identified as a co-receptor for signals that prevent axon regeneration in white matter, is by itself a negative regulator of myelination[6] and thus an inhibitor of remyelination in the adult brain. Finally, hyaluronan — a component of the extracellular matrix — accumulates around demyelinated axons to inhibit remyelination[7]. Why these molecules are expressed despite their detrimental effects is unknown. Perhaps they provide a shield around the axons to protect them from contact with activated immune cells in the brain.

Clinicians would be intrigued to know whether these variations in epigenetic regulation also affect the repair of lesions in patients with multiple sclerosis; and whether the gene expression patterns observed in the OP cells of young and old mice respectively occur in multiple sclerosis plaques that undergo remyelination or fail to remyelinate. Shen and colleagues' model of demyelination — unlike many multiple sclerosis lesions — generally shows relatively rapid and complete remyelination. Another difference is that this experimental model lacks an inflammatory milieu, which has been suggested to stimulate OP-cell differentiation and myelination[8]. These pioneering observations should therefore be repeated in other animal models of multiple sclerosis that show poor remyelination. ■

Klaus-Armin Nave is at the Max Planck Institute for Experimental Medicine, Hermann-Rein-Strasse 3, D-37075 Göttingen, Germany.
e-mail: nave@em.mpg.de

1. Shen, S. *et al. Nature Neurosci.* **11,** 1024–1034 (2008).
2. Trapp, B. D. & Nave, K.-A. *Annu. Rev. Neurosci.* **31,** 247–269 (2008).
3. Blakemore, W. F. & Franklin, R. J. *Curr. Top. Microbiol. Immunol.* **318,** 193–212 (2008).
4. Shen, S., Li, J. & Casaccia-Bonnefil, P. *J. Cell Biol.* **169,** 577–589 (2005).
5. Charles, P. *et al. Proc. Natl Acad. Sci. USA* **97,** 7585–7590 (2000).
6. Mi, S. *et al. Nature Neurosci.* **8,** 745–751 (2005).
7. Back, S. A. *et al. Nature Med.* **11,** 966–972 (2005).
8. Setzu, A. *et al. Glia* **54,** 297–303 (2006).

CANCER

Entangled pathways

René Bernards

A medley of molecules, and the interactions between them, mediate cancer. The latest news is that the enzyme CDK8 orchestrates cross-talk between two signalling pathways that are frequently deregulated in human cancers.

There is no simple answer to the question of which cellular proteins or signalling pathways are responsible for making a cell cancerous. For example, the Wnt signalling pathway, which normally plays a pivotal part in development, is often deregulated by mutation in cancer cells[1]. Mutations in the gene for the retinoblastoma tumour-suppressor protein, which is part of another signalling pathway, are also frequently associated with cancer. When put together, the findings of Firestein *et al.*[2] and Morris *et al.*[3], presented in this issue, lead to the identification of a point at which these important pathways could communicate with each other in colorectal cancer.

When Wnt proteins bind to their cell-surface receptors, this leads to the stabilization of a protein called β-catenin in the cytoplasm. β-Catenin then moves to the nucleus, where it interacts with gene transcription factors called TCFs to activate a specific set of genes. Firestein and colleagues[2] (page 547) set out to search for genes whose suppression alters β-catenin-dependent gene transcription in colon cancer cells. Of some 1,000 genes (mostly encoding kinase enzymes) that they screened, 34 seemed to be required for β-catenin activity. They then tested the same 1,000 genes for a role in the proliferation of colon cancer cells, and came up with a second list of 166 genes. The two 'hit lists' share nine genes — required for both β-catenin activity and proliferation of

DNA sequence itself. They include changes in chromatin — the complexes of DNA with histone proteins. For example, chromatin remodelling involves histone deacetylase (HDAC) enzymes, which remove acetyl groups from the lysine amino-acid residues in histone proteins. The deacetylated histones favour a more condensed chromatin conformation, which is responsible for inactivation of transcription. Also, the deacetylated histones themselves may have a reduced affinity for some transcription factors.

Shen and colleagues' previous work[4] established that, during normal development, myelination requires a reduced activity of gene transcription factors that inhibit oligodendrocyte differentiation and of those that cause OP cells to retain their stem-cell character. They also found that HDACs are essential for the differentiation of OP cells, because inhibiting these enzymes with valproic acid dramatically interfered with normal development. The authors then looked for a link between the apparent decline in remyelination with age and changes in the gene expression patterns of OP cells in adult animals.

To do this, they now examine[1] a well-established mouse model of demyelination that is generated by giving animals the toxin cuprizone. In cuprizone-treated young adult mice (eight weeks old), demyelination is followed by spontaneous remyelination. The authors find that remyelination was associated with a complex pattern of changes in gene expression that was similar to that observed during myelination in newborns. When the mice were given the HDAC inhibitor valproic acid while on a cuprizone diet, the differentiation of OP cells, and thus remyelination of white matter, was markedly impaired. Valproic acid had no obvious effect on other cell types, or on the severity of inflammation or on axon survival.

Cuprizone induced a similar pattern of demyelination in older (10-month-old) mice, but the subsequent remyelination was not accompanied by the 'molecular signature' detected in young mice. Therefore, impaired remyelination was associated with an age-dependent loss of 'epigenetic memory'. This was strikingly similar to the results seen in younger mice treated with both valproic acid and cuprizone, in particular with respect to a reduction in the expression of HDAC1 and HDAC8 by OP cells.

A closer look revealed that the impaired remyelination in older mice did not result from various possible effects — an altered inflammatory response, for example, or an abnormal increase in the number of 'reactive' astrocytes (a known side effect of brain injury), or axonal degeneration at the lesion site. Rather, genes encoding inhibitory transcription factors (such as Hes5) and factors that promote stem-cell activity (such as Sox2) were transcribed for much longer than normally seen in remyelination. Moreover, the enzyme RNA Pol II, which mediates transcription, was physically associated with these genetic loci. Thus, the ability to recruit HDACs to the promoter regions of inhibitory transcription factors, which is seen in young mice, seems to be lost in older animals.

To identify the HDACs that shut down the various inhibitors of differentiation, Shen and colleagues[1] performed gene-silencing experiments in oligodendrocytes grown in culture. The exact molecular code revealed will be of interest only to the specialist, but the message will be of great interest to stem-cell biologists and researchers in the fields of ageing and myelin-related disease: age-dependent differences in the brain's ability to repair a lesion are due to variations in the epigenetic regulation of gene expression in OP cells.

The proposed link between failing HDAC activity in the OP cells of older mice and inefficient remyelination hint at the need for 'gain-of-function' experiments. If altered HDAC activity is indeed responsible for poor remyelination, transgenic overexpression of HDAC1 in OP cells should overcome this.

But it would be too simplistic to assume that loss of epigenetic memory is the only mechanism behind poor remyelination. Various proteins also seem to directly inhibit remyelination by oligodendrocytes. For example, the cell-adhesion protein PSA-NCAM is expressed in demyelinated axons[5]. Also, LINGO-1, which was originally identified as a co-receptor for signals that prevent axon regeneration in white matter, is by itself a negative regulator of myelination[6] and thus an inhibitor of remyelination in the adult brain. Finally, hyaluronan — a component of the extracellular matrix — accumulates around demyelinated axons to inhibit remyelination[7]. Why these molecules are expressed despite their detrimental effects is unknown. Perhaps they provide a shield around the axons to protect them from contact with activated immune cells in the brain.

Clinicians would be intrigued to know whether these variations in epigenetic regulation also affect the repair of lesions in patients with multiple sclerosis; and whether the gene expression patterns observed in the OP cells of young and old mice respectively occur in multiple sclerosis plaques that undergo remyelination or fail to remyelinate. Shen and colleagues' model of demyelination — unlike many multiple sclerosis lesions — generally shows relatively rapid and complete remyelination. Another difference is that this experimental model lacks an inflammatory milieu, which has been suggested to stimulate OP-cell differentiation and myelination[8]. These pioneering observations should therefore be repeated in other animal models of multiple sclerosis that show poor remyelination. ■

Klaus-Armin Nave is at the Max Planck Institute for Experimental Medicine, Hermann-Rein-Strasse 3, D-37075 Göttingen, Germany.
e-mail: nave@em.mpg.de

1. Shen, S. *et al. Nature Neurosci.* **11,** 1024–1034 (2008).
2. Trapp, B. D. & Nave, K.-A. *Annu. Rev. Neurosci.* **31,** 247–269 (2008).
3. Blakemore, W. F. & Franklin, R. J. *Curr. Top. Microbiol. Immunol.* **318,** 193–212 (2008).
4. Shen, S., Li, J. & Casaccia-Bonnefil, P. *J. Cell Biol.* **169,** 577–589 (2005).
5. Charles, P. *et al. Proc. Natl Acad. Sci. USA* **97,** 7585–7590 (2000).
6. Mi, S. *et al. Nature Neurosci.* **8,** 745–751 (2005).
7. Back, S. A. *et al. Nature Med.* **11,** 966–972 (2005).
8. Setzu, A. *et al. Glia* **54,** 297–303 (2006).

CANCER

Entangled pathways

René Bernards

A medley of molecules, and the interactions between them, mediate cancer. The latest news is that the enzyme CDK8 orchestrates cross-talk between two signalling pathways that are frequently deregulated in human cancers.

There is no simple answer to the question of which cellular proteins or signalling pathways are responsible for making a cell cancerous. For example, the Wnt signalling pathway, which normally plays a pivotal part in development, is often deregulated by mutation in cancer cells[1]. Mutations in the gene for the retinoblastoma tumour-suppressor protein, which is part of another signalling pathway, are also frequently associated with cancer. When put together, the findings of Firestein *et al.*[2] and Morris *et al.*[3], presented in this issue, lead to the identification of a point at which these important pathways could communicate with each other in colorectal cancer.

When Wnt proteins bind to their cell-surface receptors, this leads to the stabilization of a protein called β-catenin in the cytoplasm. β-Catenin then moves to the nucleus, where it interacts with gene transcription factors called TCFs to activate a specific set of genes. Firestein and colleagues[2] (page 547) set out to search for genes whose suppression alters β-catenin-dependent gene transcription in colon cancer cells. Of some 1,000 genes (mostly encoding kinase enzymes) that they screened, 34 seemed to be required for β-catenin activity. They then tested the same 1,000 genes for a role in the proliferation of colon cancer cells, and came up with a second list of 166 genes. The two 'hit lists' share nine genes — required for both β-catenin activity and proliferation of

colorectal cancer cells. But which of these, if any, are relevant to the biology of human colon cancer?

To address this question, Firestein *et al.* investigated whether any of the nine genes resides in a region of the human genome that is amplified (meaning that additional copies of the gene are present) in colon cancer. Only one — *CDK8* — was located in a frequently amplified region, suggesting that its increased level of expression contributes to the development of colorectal tumours. Indeed, increased expression of this gene turned normal mouse fibroblast cells cancerous, and this ability depended on the presence of functional β-catenin–TCF. This observation supported the results of the authors' initial screen, which indicated that *CDK8* contributes to oncogenesis by stimulating β-catenin/TCF activity. But how does it do this?

CDK8 is a member of a family of protein kinases that must associate with regulatory proteins called cyclins to become active. CDK8 and its regulatory subunit cyclin C are components of the multi-protein Mediator complex, which couples the actions of transcription factors with the molecular machinery that carries out transcription.

When Firestein *et al.* studied the promoter sequence for one of the best-characterized target genes of β-catenin–TCF — the *c-MYC* oncogene — they found that CDK8 also interacts with this promoter close to where β-catenin–TCF binds. Decreasing the level of CDK8 by the technique of RNA interference also reduced β-catenin binding to the promoter, indicating that CDK8 (presumably as part of the Mediator complex) stabilizes the interaction of β-catenin with the *c-MYC* promoter, thereby facilitating transcription of the gene (Fig. 1a). So, amplification and overexpression of *CDK8* contributes to colon carcinogenesis, at least in part, by aiding β-catenin–TCF activity.

How does the retinoblastoma tumour-suppressor protein (pRB) fit into the picture? This protein exerts its tumour-suppressive effects mainly by restraining the activity of three closely related transcription factors of the E2F family, which regulate the expression of genes involved in both cellular proliferation and programmed cell death (apoptosis)[4]. Working with the fruitfly *Drosophila*, Morris and colleagues[3] (page 552) find that E2F1 counteracts the activity of β-catenin.

The authors show that high levels of fly β-catenin negate the apoptotic effects of increased *Drosophila E2F1* expression in the developing wing. Moreover, when *Drosophila* E2F1 was expressed in a mutant fly that expresses high β-catenin levels in the eye, the abnormal effects of the high β-catenin were greatly suppressed. What's more, E2F1 also suppressed activation of transcription by β-catenin–TCF in human cells[3]. At least in part, E2F1 exerts this effect by activating three genes (*AXIN1*, *AXIN2* and *SIAH1*), which are involved in the pathway leading to β-catenin degradation. Together, these results establish E2F1 as a negative regulator of β-catenin–TCF activity.

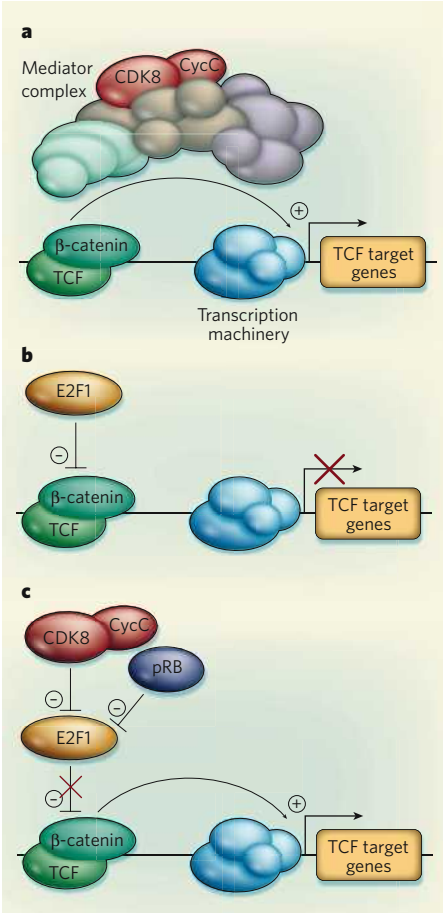


Figure 1 | Cross-talk between signalling pathways in colorectal cancer[2,3]. **a**, The product of the *CDK8* gene, which is amplified in half of the cases of colorectal cancer, stimulates transcriptional activation of β-catenin–TCF target genes, possibly as part of the Mediator multi-protein complex. CycC, cyclin C. **b**, E2F1, in contrast, inhibits the activity of the β-catenin–TCF complex. **c**, In colorectal cancer, E2F1 is itself inhibited both through phosphorylation by the CycC–CDK8 complex and by direct binding of the retinoblastoma protein pRB. Increased expression of both *RB1* and *CDK8* therefore promotes colon cancer by inhibiting E2F1 and stimulating oncogenic transcriptional activity of β-catenin–TCF.

Given the potent effects of E2F1 on β-catenin, Morris *et al.* performed a genetic screen in *Drosophila* to identify other genes that control E2F1 activity. They generated flies that have decreased expression of *Drosophila E2F1* — and so have an aberrant eye morphology — and looked for genes that could overcome the effect of this mutation. They identified one mutant that had a partial loss of *CDK8* function. This finding suggests that, in the fly, CDK8 normally acts to suppress the activity of E2F1. Indeed, expression of E2F1-regulated genes is increased in *Drosophila* larvae with mutations in the genes for either CDK8 or cyclin C (ref. 3). The observations that, in both *Drosophila* and human cells, CDK8 forms a complex with E2F1 and phosphorylates it, points to a potential mechanism for the effect of CDK8 on E2F1.

The combined results from the two studies[2,3] point to a dual effect of CDK8 on β-catenin activation. As well as the direct stimulatory effect of CDK8 on β-catenin's transcriptional activity that Firestein and colleagues[2] describe (Fig. 1a), CDK8 seems to have a second (less direct) stimulatory effect on β-catenin, by protecting it from inhibition by E2F1 (Fig. 1b). This interplay between the Wnt/β-catenin and the pRB/E2F1 signalling pathways might be even more complex, as a previous study[5] showed that cyclin D1, which when complexed with the kinase CDK4 can inactivate pRB, is a transcriptional target of β-catenin–TCF.

Together, these results[2,3] make a strong case that, at least in colon cancer, E2F1 activity must be suppressed in order to fully unleash the oncogenic activity of β-catenin–TCF. As pRB, like CDK8, is a potent inhibitor of E2F1, this means that, in the context of colorectal cancer, pRB is more likely to be acting as an oncoprotein than as a tumour suppressor (Fig. 1b). Indeed, in support of this counter-intuitive notion, although the gene encoding pRB (*RB1*) is mutated and inactivated in some 30% of all human cancers, its mutation in colorectal cancer has not been identified[6,7]. On the contrary, *RB1* is often overexpressed and even amplified in colorectal cancer[2,8]. Moreover, reduced expression of pRB inhibits proliferation of human colon cancer cells, consistent with the growth-promoting activity of this protein[3,9].

Dual oncogenic and tumour-suppressor activity has also been previously reported[10] for *E2F1*, as mice lacking this gene show tissue-specific proliferation defects as well as oncogenesis. Given the complexity of the networks mediating cancer, with new entanglements continually being revealed, there is a compelling case for the generation of a comprehensive map of genetic interactions in the human cancer genome. Undoubtedly, much more can be gleaned from studying the cross-talk between signalling pathways. ■

René Bernards is in the Division of Molecular Carcinogenesis, Center for Biomedical Genetics and Cancer Genomics Center, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, the Netherlands.
e-mail: r.bernards@nki.nl

1. Clevers, H. *Cell* **127,** 469–480 (2006).
2. Firestein, R. *et al. Nature* **455,** 547–551 (2008).
3. Morris, E. J. *et al. Nature* **455,** 552–556 (2008).
4. Rowland, B. D. & Bernards, R. *Cell* **127,** 871–874 (2006).
5. Tetsu, O. & McCormick, F. *Nature* **398,** 422–426 (1999).
6. Horowitz, J. M. *et al. Proc. Natl Acad. Sci. USA* **87,** 2775–2779 (1990).
7. Hildebrandt, B. *et al. Oncology* **59,** 344–346 (2000).
8. Gope, R. *et al. J. Natl Cancer Inst.* **82,** 310–314 (1990).
9. Williams, J. P. *et al. Mol. Cell Biol.* **26,** 1170–1182 (2006).
10. Yamasaki, L. *et al. Cell* **85,** 537–548 (1996).

Q&A

MICROBIOLOGY

Metagenomics

Philip Hugenholtz and Gene W. Tyson

Ten years after the term metagenomics was coined, the approach continues to gather momentum. This culture-independent, molecular way of analysing environmental samples of cohabiting microbial populations has opened up fresh perspectives on microbiology.

Why the 'meta' in metagenomics?

Genomics determines the complete genetic complement of an organism by high-throughput sequencing of the base pairs of its DNA. The most prominent example was the Human Genome Project, which involved the sequencing of 3 billion base pairs. But the genomes of hundreds of organisms from all three domains of life (archaea, bacteria and eukarya), as well as those of quasi-life forms such as viruses, have now been sequenced. Metagenomics, by contrast, involves sampling the genome sequences of a community of organisms inhabiting a common environment. Metagenomics has also been more broadly defined as any type of analysis of DNA obtained directly from the environment — for example, after the appropriate procedures, screening such DNA for particular enzymatic activity. To date, the approach has been applied exclusively to microbial communities.

Why do we need metagenomics?

Microbiology has traditionally been based on pure cultures grown in the laboratory. But most microorganisms cannot be grown in this way and we have been ignorant of their existence. This cultivation bottleneck has skewed our view of microbial diversity and limited our appreciation of the microbial world. Metagenomics provides a relatively unbiased view not only of the community structure (species richness and distribution) but also of the functional (metabolic) potential of a community.

What environments can be analysed?

In principle, any environment is amenable to metagenomic analysis provided that nucleic acids can be extracted from sample material (Fig. 1). Simpler communities are more tractable to a particular technique called shotgun sequencing (Box 1, overleaf) — this was a rationale for one of the earliest studies, which targeted a biofilm in acid drainage from mines that consisted of only a handful of dominant microbial populations. Most interest, however, has centred on the marine environment: the largest metagenomic study to date is the Global Ocean Sampling Expedition, which follows the voyage of Darwin's ship HMS *Beagle*. Metagenomics is now also being adopted in medicine. Of particular note is an international initiative, the Human Microbiome Project, which aims to map human-associated microbial communities (including those of the gut, mouth, skin and vagina).

What surprises have there been?

A strength of metagenomics is its potential for serendipitous discovery. An example is the discovery of proteorhodopsin proteins, light-driven proton pumps that were first identified in environmental DNA from bacterioplankton. Proteorhodopsins have since been found to be widely distributed and highly expressed in diverse microbial groups from aquatic habitats, and they may represent a major source of energy flux in the photic zone of the world's oceans. A more recent discovery is that of archaeal ammonia oxidizers. It was thought that bacteria were solely responsible for aerobic ammonia oxidation, although their numbers often could not account for the observed rates of ammonia oxidation in many habitats. The fortuitous discovery of

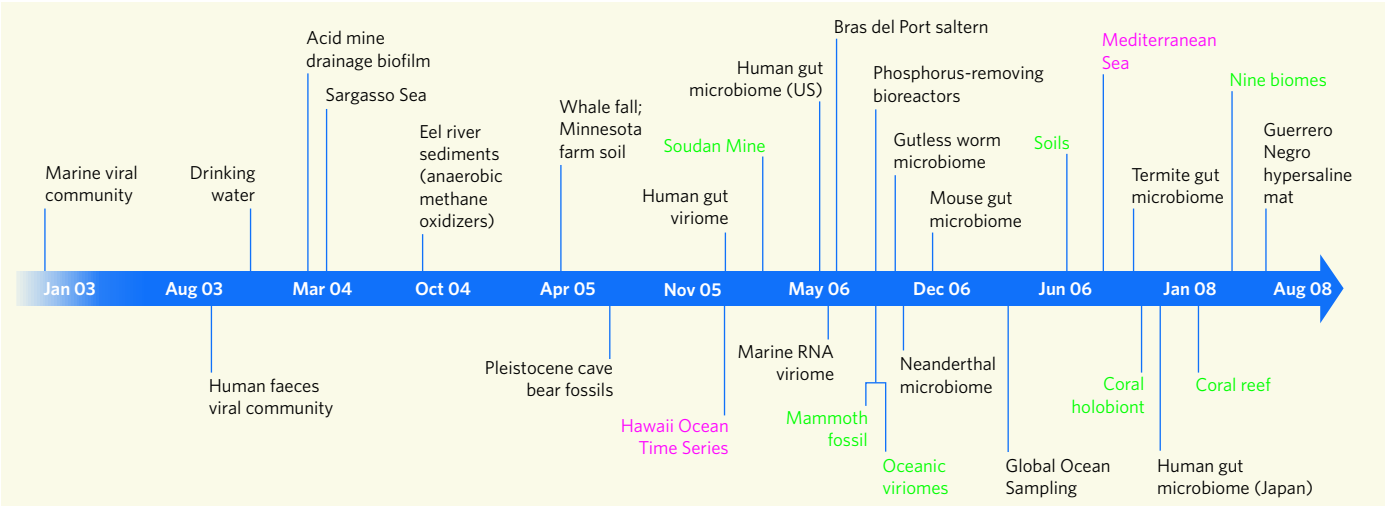


Figure 1 | Timeline of sequence-based metagenomic projects showing the variety of environments sampled since 2002. The oceanic viriomes (all viruses in a habitat) (August 2006) were from the Sargasso Sea, Gulf of Mexico, coastal British Columbia and the Arctic Ocean. The nine biomes (March 2008) were stromatolites, fish gut, fish ponds, mosquito viriome, human-lung viriome, chicken gut, bovine gut and marine viriome. The different technologies used are dye-terminator shotgun sequencing (black), fosmid library sequencing (pink) and pyrosequencing (green). (Graphic based on data sets represented at www.genomesonline.org.)

an ammonia monooxygenase gene next to an archaeal marker gene (encoding small-subunit ribosomal RNA) spawned a rash of papers implicating archaea as the main source of ammonia oxidation in many marine and terrestrial ecosystems.

Can whole genomes be reconstructed from an environmental sample?

Yes: the genomes of dominant species can be fully reconstructed from environmental samples using random sequencing. For example, complete or near-complete genomes have been assembled from microbial populations present in biofilms in acid mine drainage, in activated sludges and in marine samples. Having the complete or near-complete genome of a dominant population provides the gene inventory for the organism and allows its metabolic potential to be determined, including inferring the absence of metabolic pathways as well as their presence. A key feature of genomes obtained from environmental sources is that they are composites of the population from which they were derived, and encompass the genetic microheterogeneity present in that population.

What have we learned about microbial evolution?

Metagenomics provides the first broad insights into coexisting (sympatric) populations, as every sequence read is derived from a different individual within a given community. In communities in which deep sequence-read coverage of individual populations is possible, metagenomics provides an exquisite view of the evolutionary processes shaping these organisms. For instance, data from archaeal populations in acid mine drainage were used to show that genetic recombination occurs at a much higher frequency than previously predicted, and is the primary evolutionary force shaping these populations. And data from the Pacific and Atlantic oceans revealed that the greatest variation within populations of *Prochlorococcus* (the most abundant photosynthetic organisms in the ocean) occurs in genomic 'islands'. These islands are discrete regions in the genome that are believed to be hotspots for genomic innovation and derived in part by genes laterally transferred by viruses. Potentially the most important finding relates to the controversial topic of defining microbial species. Several metagenomic studies have shown clear discontinuity of microbial diversity, suggesting that some microbial species at least are defined by discrete 'sequence space'.

Are some environments too complex for metagenomics?

Certainly it seems unfeasible (at least for now) to obtain complete genomes of organisms from habitats with complex microbial communities because of the sheer amount of sequencing required. For example, it was estimated that a minimum of 6 billion base pairs would be required to obtain the genome sequence of the most dominant population in a soil sample, and many times that to obtain genomes from less dominant populations. But it is possible to extract biologically meaningful information from completely unassembled sequence data using a gene-centric approach.

... and the gene-centric approach is?

Rather than considering a community from the point of view of its component organisms (genomes), gene-centric analysis considers a community from the viewpoint of its component genes. Genes that are found more frequently in one community than another are assumed to endow a beneficial function on that community (Fig. 2). For example, proteorhodopsins are very common in marine surface waters compared with other habitats, and enzymes that break down cellulose are more common in the termite hindgut than in other habitats. Over-represented genes of unknown function (the majority) can provide foci for research that improves the odds of making biological discoveries. Gene-centric analysis can also be taken up a level to see if over-represented genes are part of higher functional units such as metabolic pathways. For example, genes involved in photosynthesis are generally over-represented in ocean surface waters compared with other habitats. Conversely, however, genes of low relative abundance supplying essential functions will not be detected by this approach.

Are we starting to get a handle on genetic diversity in the environment?

Scarcely: we are still far from measuring the full extent of genetic diversity encoded by microbial life. The Global Ocean Sampling Expedition has generated the largest metagenomic data set so far, comprising 6.12 million predicted proteins from 7.7 million shotgun sequences. The predicted proteins from this data set represented all previously known families of microbial proteins and added 1,700 new ones (each with more than 20 representatives). However, the rate of discovery of new protein families, containing two or more representatives, was linear with the addition of new sequences. This is not surprising in light of recent estimates of microbial diversity obtained using marker genes, such as those encoding ribosomal RNAs, which project the number of microbial species globally to be in the hundreds of millions to billions.

Can metagenomics go beyond functional prediction?

Strictly speaking, no. Metagenomics only looks at the gene sequences that encode proteins or functional RNAs. To go beyond this DNA-based view, expressed RNA transcripts and translated proteins must be obtained from environmental samples and examined directly. This is already happening. RNAs can be reverse-transcribed into DNA and sequenced, and the mass spectra of protein fragments can

Box 1 | The nuts and bolts of metagenomics

Metagenomics begins with the extraction of genomic DNA from cellular organisms and/or viruses in an environmental sample. For 'traditional' dye-terminator sequencing, the DNA is sheared into uniform lengths and inserted into a vector — a known DNA fragment that can be moved between organisms. The vector is then replicated in a bacterial host (typically *Escherichia coli*), which produces many clones of the genome fragment suitable for sequencing. Current sequencing technologies produce 'reads' of fewer than 1,000 bases (Fig. 3), so thousands of reads are required to re-cover whole genomes. As with single-genome (isolate) projects, a range of sizes of DNA can be cloned.

Initially, metagenomic studies focused on screening libraries of large-insert clones (fosmids and bacterial artificial chromosomes) for clones containing genes of functional or evolutionary interest, and these clones were then fully sequenced to provide contextual data. This is still a useful approach, but such directed sequencing has largely been replaced by 'shotgun' sequencing of randomly sampled, anonymous DNA, which is cheaper and has higher throughput.

As with isolate genomics, metagenomic data-processing usually involves the assembly of short, overlapping sequence reads into a consensus sequence, and prediction of which stretches of sequence encode genes. In metagenomics, however, complex communities may not result in any assembly, as no population may have been sampled a sufficient number of times to result in overlapping reads.

The amount of an organism's genome recovered from an environmental sample depends on how many sequence data are obtained and the relative abundance of the organism in its community. For example, assuming completely random sampling, a population with an average genome size of 3 million base pairs, and comprising 0.1% of a community, would require 3 billion bases of sequence data to obtain a 1× coverage (each base of the genome is represented on average by one read). This is beyond the range of dye-terminator sequencing, but new, highly parallelized sequencing technologies, such as 454–Roche pyrosequencing and Illumina sequencing, may be up to the job (Fig. 3).

A desirable extra step in metagenomics is to identify the owner of each anonymous DNA fragment, a process called binning, or classification. Binning methods are still in their infancy, and usually only longer fragments (5 kilobases or more) can be classified reliably. **P.H. & G.W.T.**

be determined and identified by reference to their gene sequences. These approaches have given rise to the fields of metatranscriptomics and metaproteomics.

Can metagenomics be used for viruses?

Certainly. Viral communities have been the subject of several metagenomic investigations and were among the earliest to be studied using the method (Fig. 1). These studies, and viral sequences cropping up in metagenomic data in general, all point to a central role for viruses in microbial evolution and ecology. Initially, only double-stranded DNA viruses were accessible through cloning, but the newer cloneless sequencing technologies allow access to all types (such as single-stranded and RNA viruses).

... and for eukaryotes?

Yes. But because of the cost, sequencing projects have tried to avoid them: eukaryotes in general have much larger genomes and a higher proportion of DNA that doesn't code for protein. Metagenomic analyses of insect–microbe symbioses, for example, have tried to exclude the eukaryotic host DNA. As sequencing costs continue to fall, particularly with the development of higher-throughput technologies, eukaryotes should become a tractable component of a metagenomic analysis.

What will eukaryotic metagenomics tell us?

A common misconception is that animals and plants constitute the bulk of eukaryotic diversity. In fact, most of the evolutionary diversity exists in the under-studied microbial eukaryotes, which are typically unicellular and as difficult to culture as their bacterial and archaeal cousins. For this reason, metagenomics will be an excellent way to study microbial eukaryotic biology (and thereby eukaryotic biology as a whole). However, this has not yet happened because most microbial eukaryotes also have large genomes. Even for plants and animals, a metagenomic approach may be warranted. Basing our understanding of a species on the genome sequence of one or a few individuals misses a lot of genetic information, as is beginning to be appreciated for the human population. In the case of large multicellular eukaryotes such as humans, the equivalent of metagenomics is to sequence the genomes of many individuals.

Will computational capacity keep pace?

That remains to be seen. Sequence data are increasing at a rate higher than increases in computational power. Even more problematic than simple storage of sequence data are the 'all-versus-all' sequence comparisons required to best interpret metagenomes, which raise the computational requirements exponentially. Unless there is a radical breakthrough in computing, for example if quantum computers become viable, then all-versus-all comparisons of metagenomic data will not be feasible in the near future. On the upside, it is unnecessary to compare all data with each other to extract biological insights, and the biology can drive the selection of a relevant subset of available metagenomic data.

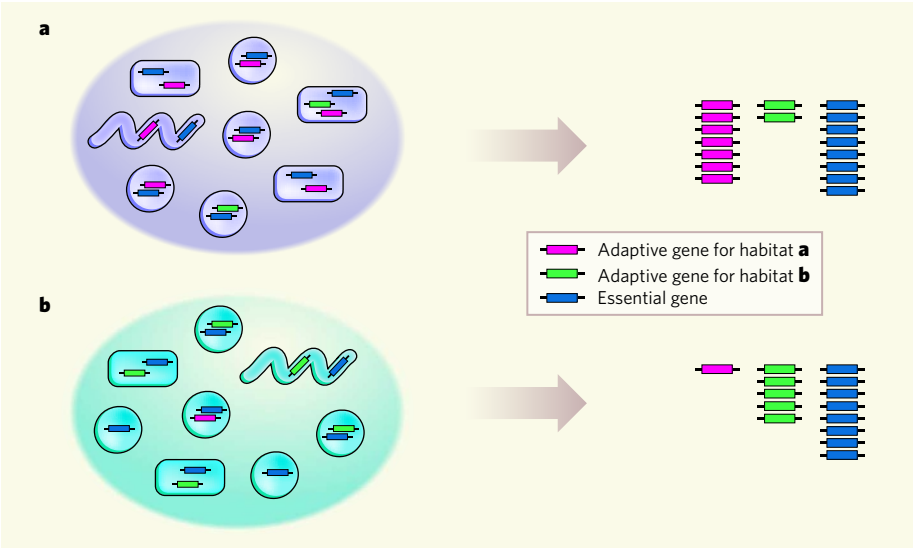


Figure 2 | Gene-centric analysis. Genes from communities **a** and **b** are assigned to their respective gene families (red, green or blue) and counted to highlight functions that are putatively beneficial (red over-represented in community **a**, and green over-represented in community **b**), or essential or having a generalized function (blue; high counts in both communities). The organisms that contribute the genes are largely ignored in this approach, and the underlying assumption is that high counts indicate functions that are important for survival in a given habitat. (Graphic by S. Tringe.)

What other bottlenecks are there?

The gap between characterized and hypothetical proteins identified in metagenomes is widening at an alarming rate. Next to computational resources, uncharacterized gene products are likely to be the biggest bottleneck for the foreseeable future. This means that our understanding of microbial ecosystems will be partial at best, being based on what we can infer from our existing knowledge of biochemistry. Before the advent of metagenomics, however, we were completely oblivious to what we didn't know (unknown unknowns); now we have the blueprints, although we can't read many of the instructions (known unknowns).

Platform	Million base pairs per run	Cost per base (US¢)	Average read length (base pairs)
Dye-terminator (ABI 3730xl)	0.07	0.1	700
454–Roche pyrosequencing (GS FLX titanium)	400	0.003	400
Illumina sequencing (GAii)	2,000	0.0007	35

Figure 3 | Comparison of the cost and throughput of sequencing technologies. New technologies (454–Roche pyrosequencing and Illumina sequencing) generate far more sequence data per run, at a much lower cost than conventional dye-terminator sequencing, but the reads are shorter. Improvements in these technologies are already producing longer reads, and single-molecule sequencing (an as-yet-unproven emerging technology) holds the promise of longer reads still. The dye-terminator approach may soon be obsolete.

Who can use metagenomics?

An increasing number of labs: metagenomics is becoming a basic technology for understanding the ecology and evolution of microbial ecosystems, upon which hypotheses and experimental strategies are built. And with new sequencing technologies producing more than 100 megabases of data for less than US$20,000 (Fig. 3), metagenomics is now within the reach of many researchers. ■

Philip Hugenholtz is in the Microbial Ecology Program, DOE Joint Genome Institute, Walnut Creek, California 94598, USA.
Gene W. Tyson is in the Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.
e-mails: phugenholtz@lbl.gov; gtyson@mit.edu

FURTHER READING

Tyson, G. W. *et al.* Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* **428,** 37–43 (2004).
Edwards, R. A. & Rohwer, F. Viral metagenomics. *Nature Rev. Microbiol.* **3,** 504–510 (2005).
DeLong, E. F. *et al.* Community genomics among stratified microbial assemblages in the ocean's interior. *Science* **311,** 496–503 (2006).
Turnbaugh, P. J. *et al.* An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444,** 1027–1031 (2006).
Yooseph, S. *et al.* The *Sorcerer II* Global Ocean Sampling Expedition: expanding the universe of protein families. *PLoS Biol.* **5,** e16 (2007).
Warnecke, F. *et al.* Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* **450,** 560–565 (2007).

ARTICLES

Frequency-modulated nuclear localization bursts coordinate gene regulation

Long Cai[1]*, Chiraj K. Dalal[1]* & Michael B. Elowitz[1]

In yeast, the transcription factor Crz1 is dephosphorylated and translocates into the nucleus in response to extracellular calcium. Here we show, using time-lapse microscopy, that Crz1 exhibits short bursts of nuclear localization (typically lasting 2 min) that occur stochastically in individual cells and propagate to the expression of downstream genes. Strikingly, calcium concentration controls the frequency, but not the duration, of localization bursts. Using an analytic model, we also show that this frequency modulation of bursts ensures proportional expression of multiple target genes across a wide dynamic range of expression levels, independent of promoter characteristics. We experimentally confirm this theory with natural and synthetic Crz1 target promoters. Another stress-response transcription factor, Msn2, exhibits similar, but largely uncorrelated, localization bursts under calcium stress suggesting that frequency-modulation regulation of localization bursts may be a general control strategy used by the cell to coordinate multi-gene responses to external signals.

Cells sense extracellular signals and respond by regulating the expression of target genes[1–3]. This process requires two stages of information processing. First, cells encode extracellular signals internally, in the states and localization of transcription factors. Second, transcription factors activate the expression of downstream genes that will implement cellular responses[1–3]. Although many signal transduction systems have been studied extensively, it often remains unclear how signals are encoded dynamically in transcription factor activities at the single-cell level. In addition, cellular responses often involve many proteins acting together, rather than individually. However, in general it is not known how the expression levels of target genes are coordinated, allowing them to be regulated together, despite diverse promoter architectures[3]. Here we investigate how signal encoding and protein coordination are achieved in individual cells.

We examined the calcium stress response pathway in *Saccharomyces cerevisiae*, or budding yeast. Cellular responses to extracellular calcium are mediated by Crz1, the calcineurin-responsive zinc finger transcription factor[4]. The activity of Crzl is modulated by phosphorylation and dephosphorylation[4], resulting in changes in the nuclear localization of Crzl protein (Fig. 1a), rather than changes in its abundance (Supplementary Fig. 1). To understand how Crzl phosphorylation dynamics respond to calcium and regulate the more than 100 different targets necessary for calcium adaptation[5], we acquired time-lapse movies of Crzl localization dynamics, using a strain in which the Crzl protein was tagged with green fluorescent protein (GFP)[6]. In each movie, we tracked the response of Crzl localization in individual cells to step changes in extracellular calcium concentration. We found that Crzl dynamics connect the encoding of signals and the coordination of target gene expression.

Frequency modulated bursts of nuclear localization

In the absence of calcium, Crz1 was cytoplasmic in all cells (Supplementary Movie 1). Upon the addition of calcium, individual cells exhibited a rapid, synchronized burst of Crz1 nuclear localization, similar to behaviour observed with the yeast osmosensor Hog1 (refs 7, 8). However, unlike Hog1, this initial burst was followed by sporadic unsynchronized localization bursts, typically lasting about 2 min (Fig. 1b–d) and persisting throughout the course of the movie (up to 10 h, see Supplementary Movie 1). Moreover, these single-cell Crz1 dynamics are consistent with microarray studies performed on cell populations[5]: after a step change in calcium, an initial overshoot in messenger RNA levels of Crz1 target genes results from the initial synchronous burst of Crz1, whereas the subsequent elevated average expression levels are due to sustained unsynchronized bursts in individual cells (Fig. 1c–e).

We next addressed how the amount of calcium affects the dynamics of nuclear localization. We observed that the fraction of cells with nuclear-localized Crz1 increased with calcium concentration. Because Crz1 localizes in bursts, this calcium dependence could in principle result from increases in burst frequency or duration. Strikingly, analysis of movies revealed that only the burst frequency increased (Fig. 2a), whereas the distribution of burst durations remained constant at all calcium concentrations (Fig. 2b). This distribution was consistent with two rate-limiting stochastic steps, each with a timescale of about 70 s (Fig. 2b). Thus, cells use stereotyped Crz1 localization bursts in a frequency-modulated fashion to encode and respond to extracellular calcium. This contrasts with amplitude-modulation control, in which the fraction of Crz1 molecules found in the nucleus would change with calcium, but remain constant over time.

Movies also revealed two modes of nuclear localization bursts: isolated individual bursts and clusters of bursts, analogous to spike trains in neurons (Fig. 1c, d). At calcium concentrations less than 100 mM, only isolated bursts were observed and the averaged autocorrelation function of the localization trajectories from individual cells was well fitted by a single exponential. However, as clustered bursts emerged at calcium concentrations greater than 100 mM (Fig. 1c, d), the averaged autocorrelation function was better fit by a sum of two exponentials, whose timescales matched the typical durations of isolated bursts and burst clusters, respectively (Fig. 2d). Higher levels of calcium led to an increasing proportion of bursts occurring in clusters (Fig. 2d inset). Eventually, at the highest calcium levels, Crz1 nuclear localization trajectories appeared more similar to sustained oscillations[9] than to the isolated stochastic bursts seen at lower calcium concentrations.

[1]Howard Hughes Medical Institute, Division of Biology and Department of Applied Physics, California Institute of Technology, M/C 114-96, Pasadena, California 91125, USA.
*These authors contributed equally to this work.

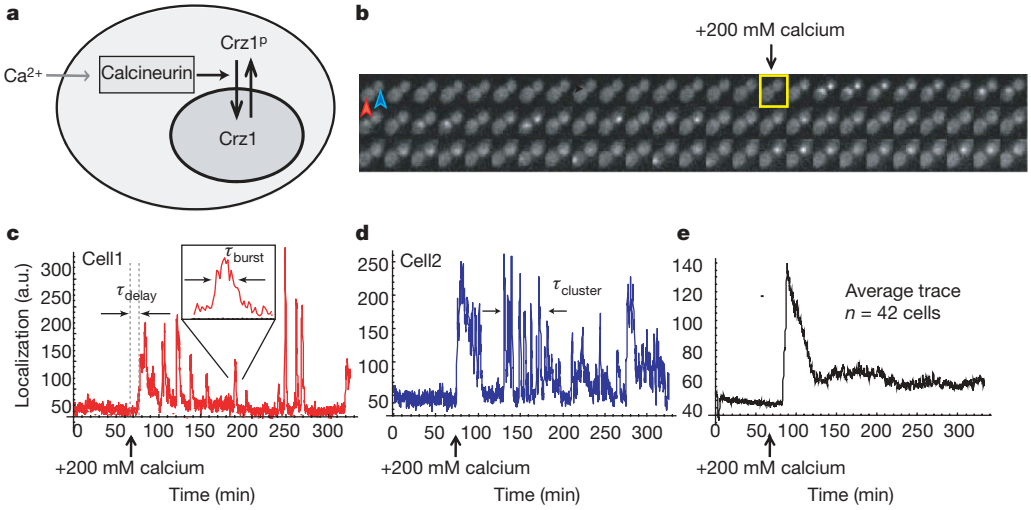


Figure 1 | Crz1 undergoes bursts of nuclear localization in response to calcium. **a**, In the presence of extracellular Ca^{2+}, Crz1 is dephosphorylated and translocates into the nucleus. **b**, Filmstrip showing yeast cells with Crz1–GFP before and after addition of 200 mM extracellular calcium (yellow square). Frames displayed here are separated by 4.5 min, but actual time resolution is higher. **c**, **d**, Two single-cell time traces showing Crz1 localization behaviour of the two cells in **b**. Note that there is a synchronized initial burst of nuclear localization followed by subsequent unsynchronized isolated and cluster bursts of localization. Individual burst duration, τ_{burst}, and cluster duration, $\tau_{cluster}$, as well as the delay between calcium addition and the initial response, τ_{delay}, are defined on the traces. **e**, Averaged localization trace shows how single-cell burst dynamics yield partial adaptation across a population of cells.

An additional level of quantization emerged in the initial response to a change in calcium, which exhibited an all-or-nothing response at the single-cell level. At calcium concentrations of at least 100 mM, most cells displayed a synchronous initial burst of Crz1 nuclear localization. This resulted in a sharp histogram of values of τ_{delay}, defined as the time interval between calcium addition and the first burst of Crz1 nuclear localization (Fig. 1c). In contrast, at lower calcium levels of about 10 mM, no synchronous initial response occurred, as reflected in the much broader distribution and larger mean of τ_{delay}. Intermediate calcium concentrations produced a mixture of the two discrete behaviours (Supplementary Fig. 2). Thus, rather than controlling the amplitude of the initial response, calcium modulates the proportion of fast-responding cells (Fig. 2c). Critically, in both the initial response and the subsequent bursts, we never observed persistent intermediate levels of localization. Taken together, these results reveal that Crz1 burst activity is 'quantized' at multiple levels in cells and that only the frequencies of burst events are modulated by extracellular calcium.

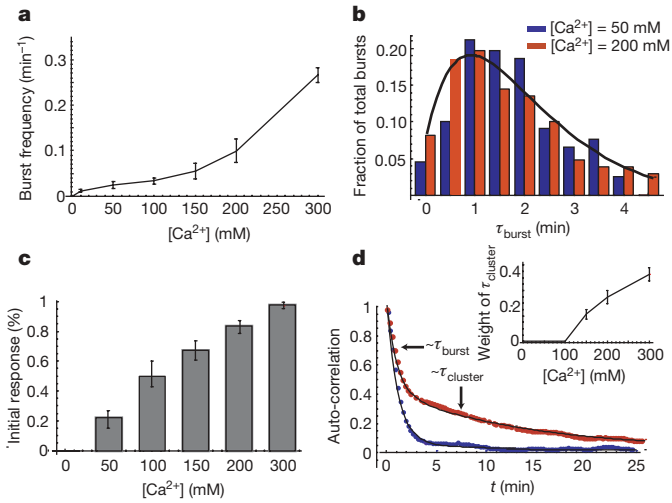


Figure 2 | Calcium modulates the frequency, but not the duration, of Crz1 nuclear localization bursts. **a**, Frequency of bursts increases with calcium concentration (error bars calculated by using different thresholds for burst determination, see Supplementary Information). **b**, Burst duration is independent of calcium concentration. Normalized histograms, $h(t)$ of total burst duration at two calcium concentrations are both well fit by $h(t) = te^{-t/\tau}$ with $\tau = 70$ s (black line). **c**, The proportion of cells that respond initially to extracellular calcium increases with the calcium concentration. Bars represent the fraction of cells with nuclear-localized Crz1 within 15 min of addition of calcium. **d**, Average autocorrelation functions of localization trajectories ($n = 58$ and 85 cells respectively) from a population of cells at two calcium concentrations. At low Ca^{2+} concentrations (blue, 50 mM), the autocorrelation can be well fit by a single exponential with timescale $\tau_{burst} \approx 60$ s, whereas at high Ca^{2+} concentrations (red, 200 mM), two timescales of fluctuations emerge, $\tau_{burst} \approx 60$ s and $\tau_{cluster} \approx 720$ s, corresponding to isolated and clustered bursts, respectively. Inset shows the relative weight of the clustered bursts, which appear at Ca^{2+} concentrations greater than 100 mM and increase in frequency as calcium increases. Error bars are estimated from bootstrap.

Active generation of bursts

Because nuclear localization bursts occurred stochastically in single cells and were unsynchronized within a population on the same slide, they could not be driven by fluctuations in external conditions. These bursts also cannot be explained simply by independent fluctuations in the phosphorylation or localization state of individual Crz1 molecules, because each burst involves coherent translocation of a large fraction of the approximate 1,000 copies of Crz1 present per cell[10]. To gain insight into these dynamics, we next asked whether Crz1 bursts were related to, or driven by, other dynamic cellular phenomena.

First, we acquired movies of cells expressing both the G1 cell-cycle phase marker Whi5–GFP[11] and Crz1–mCherry. We found no evidence for cell-cycle regulation of Crz1 localization bursts, nor were bursts in daughter cells correlated with those in corresponding mother cells (Supplementary Fig. 5).

Second, we examined the role of intracellular calcium, which in other cell types has been shown to exhibit spike-like dynamics[12]. We acquired movies of yeast cells expressing both a fluorescence resonance energy transfer (FRET)-based calcium sensor[13] and Crz1–mCherry. We observed sporadic transient spikes in intracellular calcium lasting about 38 s (Supplementary Fig. 7). These spikes coincided with some Crz1 localization bursts (Fig. 3a and Supplementary Fig. 6). However, Crz1 localization bursts also occurred without corresponding calcium spikes, indicating that localization bursts may be stimulated by calcium spikes, but are not exclusively determined by them.

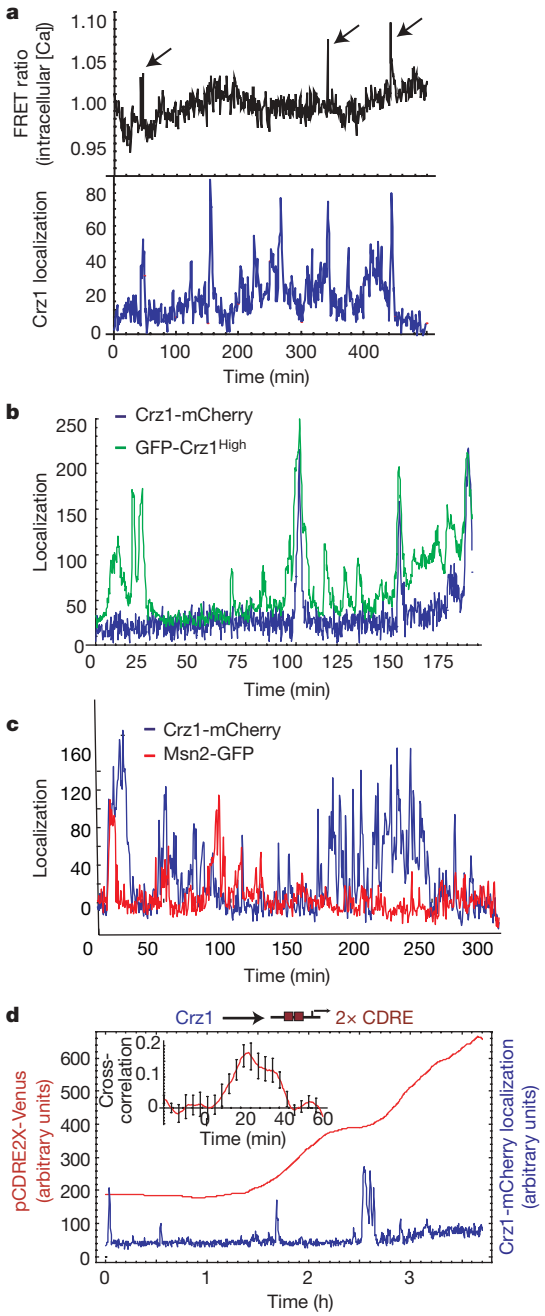


Figure 3 | Crz1 localization bursts are partially independent of other cellular processes and affect downstream gene expression. **a**, Time traces of Crz1–mCherry localization (blue) in arbitrary units and FRET ratio changes (black), indicating intracellular calcium levels. Arrows indicate spontaneous calcium spikes coincident with Crz1 localization bursts. **b**, Single-cell time traces of Crz1–mCherry (blue) and the Crz1^{high}–GFP mutant (green) with increased affinity to calcineurin. Both Crz1 proteins are expressed and measured simultaneously in the same cell. **c**, Single-cell traces of Crz1–mCherry and Msn2–GFP in the same cell. Note that the two proteins exhibit statistically similar burst-like behaviour but only weak correlation. **d**, Crz1–mCherry localization (blue) increases expression of the Crz1 target synthetic promoter (2×CDRE) (red). Transcriptional bursts in p2×CDRE–Venus are preceded by corresponding Crz1 localization bursts. Note that not all Crz1 localization bursts result in observable transcriptional bursts, suggesting additional levels of stochasticity in transcription initiation. Inset shows positive cross-correlation ($n = 9$ cells) between the promoter activity and Crz1 localization with a delay corresponding to target protein maturation. Error bars are estimated from bootstrap.

Third, we asked whether Crz1 bursts are driven by fluctuations in calcineurin, the upstream phosphatase that initiates Crz1 nuclear localization[14]. We analysed a Crz1 mutant[15] with higher affinity to calcineurin, GFP-Crz1^{high}, together with Crz1–mCherry simultaneously in the same cell. If Crz1 bursts were just passive readouts of spikes in calcineurin activity, one would expect both types of Crz1 to burst simultaneously, with amplitudes related to their relative calcineurin affinities. Instead, we observed that Crz1^{high} exhibited an increased burst frequency. Wild-type Crz1 bursts were a subset of these Crz1^{high} bursts (Fig. 3b). Similarly, sub-saturating concentrations of the calcineurin inhibitor FK506 (ref. 16) reduced the frequency of bursts, but did not affect their amplitude (Supplementary Fig. 8). As the calcineurin–Crz1 interaction controls only the frequency of burst initiation, Crz1 localization dynamics do not simply follow upstream fluctuations in calcineurin activity.

Finally, to investigate the generality of localization bursts, we examined Msn2, a general stress-response transcription factor that was previously reported to exhibit nuclear localization oscillations[17–19]. We found that Msn2–GFP localization is induced by calcium stress. Like Crz1, it localized in short bursts on a timescale of 1.5–2 min, and exhibited clustered bursts (Fig. 3c and Supplementary Fig. 9). However, despite these statistical similarities, bursts of the two proteins were largely uncorrelated when observed simultaneously in the same cells under calcium stress (Fig. 3c and Supplementary Fig. 10). Of the Msn2–GFP bursts, 18.4 ± 2.0% coincided with Crz1–mCherry bursts, a fraction only slightly higher than the 14.0 ± 1.2% of overlapping events expected by chance if the two proteins burst independently (Supplementary Information). Furthermore, bursts of Msn2 nuclear localization can occur in the absence of calcium, whereas Crz1 bursts cannot. Similarly, we observed that the glucose-responsive repressor Mig1 exhibited bursts of nuclear localization at low glucose concentrations. These results suggest that cells operate multiple transcription-factor localization burst systems in a largely independent manner.

The functional role of frequency modulation

What effect do frequency-modulation-regulated Crz1 localization bursts have on downstream genes? We analysed the transcriptional activity of a synthetic Crz1-dependent promoter containing two calcineurin-dependent response elements (2×CDRE)[20], driving expression of the fluorescent protein Venus. We also monitored Crz1–mCherry localization simultaneously in the same cell (Fig. 3d). We found that the promoter was expressed in transcriptional bursts that followed Crz1 localization bursts (Fig. 3d). Interestingly, not all Crz1 localization bursts resulted in observable transcriptional bursts, suggesting that transcription initiation is probabilistic. Nevertheless, the rate of Venus production (time derivative of fluorescence) was correlated with Crz1–mCherry bursts with a time-delay comparable to the maturation time of the Venus fluorophore, as expected (Fig. 3d inset and Supplementary Fig. 12). We observed similar results with a natural Crz1 target gene, Cmk2 (Supplementary Fig. 11)[5]. Thus, transcription-factor localization bursts propagate to downstream targets, and represent a general mechanism for generating 'transcriptional bursting'[21–23] in downstream gene expression[24–34].

Standard models of gene regulation involve amplitude modulation, or analogue, changes in transcription-factor concentration in response to external signals. What biological functions could the frequency modulation form of regulation observed here provide for the cell? Crz1 regulates more than 100 different target genes[5], including Ca^{2+} pumps and structural proteins necessary for calcium adaptation. The target promoters of these genes may differ in their input functions, defined as the dependence of transcription rate on the concentration of transcription factor in the nucleus. Input functions vary widely in their minimal and maximal levels of expression, the concentration of transcription factor at which they reach half-maximal activity, and in the sharpness, or cooperativity, of their

response[35]. We used an analytic model to show that frequency-modulation regulation of nuclear localization bursts could allow transcription factors to modulate the expression of multiple target genes in concert, keeping their relative abundances fixed over a wide dynamic range, regardless of the shapes of their input functions.

This hypothesis can be understood by comparing the effects that amplitude- and frequency-modulation regulation systems have on two hypothetical target promoters, labelled A and B, with different input functions (Fig. 4b, e, grey dashed curves). In the amplitude-modulation regulation system, the fraction of Crz1 molecules in the nucleus, $Crz1_{nuc}$, would increase with calcium, but remain constant over time (Fig. 4a, b). Accordingly, histograms of $Crz1_{nuc}$ would exhibit a single peak whose position is calcium-dependent (Fig. 4b). The expression level of each gene would be proportional to its input function evaluated at this peak position. Because it is expected that these input functions would, in general, differ in shape, the normalized rates of A and B expression would vary differently with $Crz1_{nuc}$, and hence their ratio would vary with calcium (Fig. 4c). Thus, in the amplitude modulation model, A and B would be expressed in an 'uncoordinated' fashion.

In contrast, frequency-modulation regulation (Fig. 4d) would control the fraction of time that Crz1 is nuclear, rather than the concentration of nuclear Crz1. In the limiting case of switching between two localization levels, this would cause $Crz1_{nuc}$ to be either high, during a burst, or very low, between bursts, but rarely in between, resulting in a bimodal $Crz1_{nuc}$ histogram (Fig. 4e). The expression level of each target promoter would be determined mainly by the value of its input function near the location of the higher histogram peak, and by the fraction of time Crz1 spends in the localized state, which determines the height of the peak. In frequency-modulation regulation, higher levels of calcium would increase the relative height of this peak, by increasing the frequency of bursts, but would not change its position

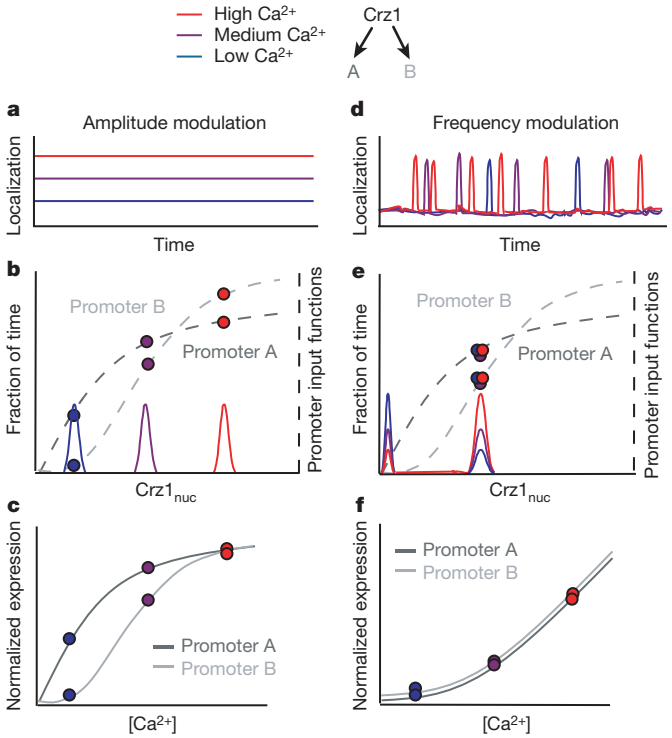


Figure 4 | Frequency- versus amplitude-modulation regulation of two hypothetical target genes, labelled A and B (schematic). **a**, In the amplitude-modulation regulation system, the fraction of nuclear Crz1 ($Crz1_{nuc}$) changes with calcium, but remains constant over time. **b**, As such, the histogram of $Crz1_{nuc}$ yields single peaks at calcium-dependent positions. Target gene expression level is proportional to the input functions at these peak positions. **c**, Because their input functions differ, the normalized rates of A and B expression vary differently with nuclear Crz1, and hence with calcium, yielding different (uncoordinated) expression profiles as a function of calcium. **d**, In frequency modulation, where Crz1 molecules collectively move between nuclear and cytoplasmic compartments, $Crz1_{nuc}$ is either high or low, during or between bursts, respectively. This graph depicts the limiting case of rapid and complete transitions between two states, but results do not depend on this assumption (see Supplementary Information). **e**, This yields a bimodal histogram in which the height, but not the position, of the peak is calcium dependent. **f**, Consequently, the expression levels of A and B are each proportional to burst frequency, and hence to each other, yielding coordinated expression, as shown.

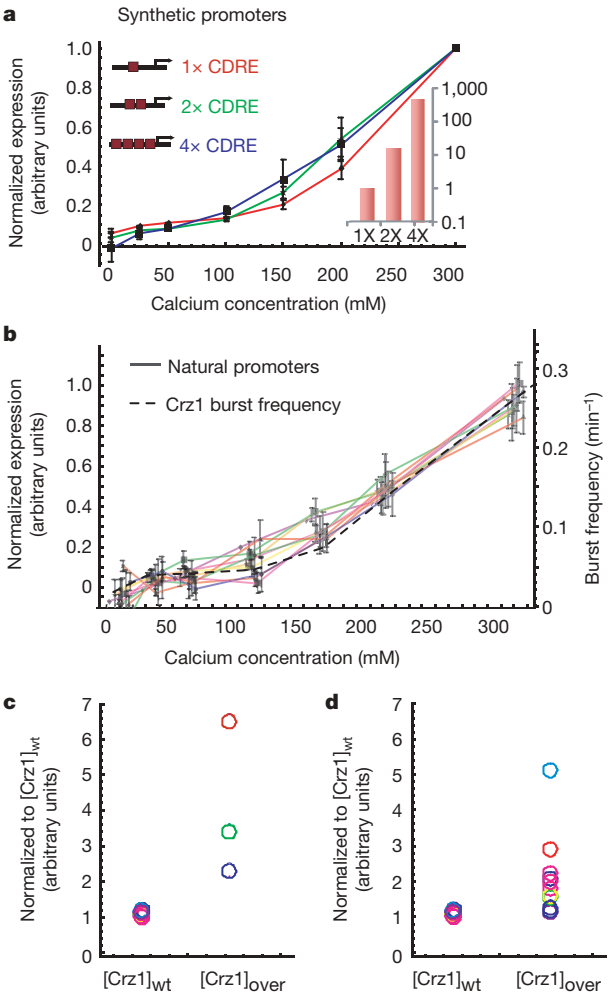


Figure 5 | Frequency-modulated bursts coordinate gene expression. **a**, Measured expression levels of synthetic Crz1 target promoters containing one, two and four CDREs. In each case, expression levels are normalized by their maximum. Note the similarity of curves to each other and to burst frequency. Bottom inset shows expression at 200 mM calcium normalized by the expression of the 1×CDRE. **b**, Expression profiles of natural target promoters (solid lines) exhibit similar inductions as the Crz1 burst frequency (dashed line). Only ten are plotted here for clarity. Error bars indicate standard error from repeated experiments. Induction curves for 64 genes are shown in Supplementary Fig. 14. **c**, Expression level of synthetic target genes at wild-type versus overexpressed levels of Crz1. Each promoter is normalized by its maximum expression in cells with wild-type Crz1 levels. Thus the normalized expression levels of one, two and four CDREs at $[Crz1]_{wt}$ each equal 1. If all promoter input functions were identical, then their normalized expression should increase by the same factor when Crz1 is overexpressed. However, the data show a range of fold changes in expression levels, excluding identical input function shapes as an explanation for proportional regulation. **d**, Same as **c** for the natural promoters in **b**.

(Fig. 4e). The expression levels of the two genes would therefore be individually proportional to the burst frequency, and consequently remain proportional to each other, as calcium is varied. Thus, A and B would be coordinated, expressed at a constant ratio at all calcium levels, regardless of the shapes of their input functions (Fig. 4f).

This regulatory strategy is general: it does not require bursts to saturate target promoter input functions, nor does it require that the $Crz1_{nuc}$ histogram be bimodal (see Supplementary Information for a more detailed treatment). Thus, it functions with bursts, such as those observed here, that span a range of localization amplitudes, giving rise to a non-bimodal histogram. Finally, this strategy is more general than frequency-modulation regulation; it can also work by varying the duration, rather than the frequency, of localization bursts. We refer to this principle as FM-coordination.

Experimental validation of FM-coordination

To test the FM-coordination hypothesis experimentally, we first analysed three synthetic promoters containing one, two or four CDREs. These promoters were previously shown to exhibit Crz1-dependent expression in response to calcium[20]. We used flow cytometry to analyse strains containing each of these three promoters fused to yellow fluorescent protein (YFP). The promoters varied in their expression strength over two orders of magnitude, in the ratio 1:15:450, respectively (Fig. 5a inset). Nevertheless, they exhibited identical dependence on calcium, collapsing onto a single curve when normalized, as predicted by FM-coordination (Fig. 5a).

Natural eukaryotic promoters exhibit more diverse architectures than these pure synthetic promoters. Therefore, we also investigated whether FM-coordination could explain the behaviour of natural Crz1 target genes. A previous microarray study identified 163 potential Crz1 target genes[5], 83 of which were available as fluorescent protein fusions[6]. Of these, 40 showed a measurable and predominantly calcineurin-dependent response to calcium under our media conditions (Supplementary Fig. 15). These natural promoters, like their synthetic counterparts, exhibited a broad distribution of expression levels (Supplementary Figs 14 and 17). Notably, a large subset of these target genes (34 out of 40) exhibited normalized calcium response curves identical to each other and to those of the synthetic promoters. They were thus regulated in proportion to burst frequency, in accordance with FM-coordination (Fig. 5b and Supplementary Fig. 14a).

In principle, target promoters could respond proportionally because they share identical input functions. If so, increasing total Crz1 concentration would increase the expression of all promoters by the same factor (Fig. 5c, d). Conversely, if the expression levels of the promoters change by different factors when Crz1 concentration is increased, then their input functions must differ in shape. This is what we observed when we compared the expression of both the synthetic and natural promoters in wild-type strains with their expression in a Crz1-overexpressing strain[36]. The increase in expression varied between onefold and sevenfold over this set of promoters (Fig. 5c, d), confirming that synthetic and natural promoters are proportionally regulated by the FM-coordination mechanism, despite non-identical input functions.

Discussion

Diverse perturbations including calcium, pharmacological inhibitors and genetic mutations vary mean Crz1 activity across a broad dynamic range exclusively by affecting the frequency of localization bursts, rather than their duration or amplitude. Thus, Crz1 activity is quantized both in its initial response to a step change in calcium, which is all-or-nothing at the single-cell level, and in its sustained response, which is composed of stereotyped bursts of localization. These data suggest that eukaryotic cells can encode information about the extracellular environment in the frequency of stochastic intracellular events, rather than in the concentrations of molecular species. This contrasts with canonical signal transduction pathways, in which concentrations of activated proteins are used to convey information to the nucleus. Using localization as a single-cell reporter for post-translational modifications, we observed qualitatively similar, but largely uncorrelated, behaviour in the Msn2 stress-response transcription factor and the Mig1 glucose repressor, suggesting that localization bursts are broadly used in the cell.

In engineering, frequency modulation appears in diverse signal processing and control applications such as broadcasting[37] and rocket thruster control[38]. Spike frequency control, or rate coding, is also fundamental to neural computation[39], where it may overcome noise background for signal propagation, among other functions[40]. In genetic networks, frequency-modulation regulation solves a fundamental problem, allowing cells to co-regulate proportionally a large set of target genes with a diverse range of input functions. Note also that these results represent a complementary mode of regulation compared with previous observations of calcium oscillation frequency-dependent gene expression[41]. FM-coordination connects the dynamic behaviour of a single protein to the expression of large regulons, and may thus provide insight into genome-scale expression patterns. Furthermore, at the evolutionary level, FM-coordination enables promoter mutations to alter the level of expression of individual genes without disrupting their coordination. In contrast, to achieve coordinate regulation using amplitude modulation would require 'fine-tuning' of target promoter input functions, severely constraining promoter architecture and regulatory evolution. In light of previous observations of transcription-factor pulsing in p53 (ref. 42), NFκB[43] and SOS stress-response systems[44], and because of its potential use in protein and metabolic, as well as transcriptional, networks, we anticipate that frequency-modulated regulation may represent a general principle by which cells coordinate their response to signals.

METHODS SUMMARY

Strain construction. All GFP strains were obtained from the GFP carboxy-terminal protein fusion library (Invitrogen)[6]. All other strains were constructed similarly to the library: polymerase chain reaction (PCR) was used to amplify fluorescent proteins[45], which were then transformed into the w303a laboratory strain of yeast[46,47]. FRET[13] and Crz1 mutant plasmids[15] were transformed into a Crz1–cherry strain. A Crz1 overexpression plasmid[46] was transformed into 83 Crz1 target gene–GFP strains. YFP was ligated into synthetic promoter (CDRE) plasmids[20] and transformed into laboratory and Crz1–cherry strains.

Time-lapse microscopy. Cells were attached to glass-bottom dishes (Willco) functionalized with concanavalin-A (Sigma). Fluorescence images were taken at room temperature on an Olympus IX81 with the ZDC autofocus option and an Andor Ikon (DU-934) camera. $CaCl_2$ was added during movies to nominal concentrations using a 2× stock solution. Images were acquired at 10-s intervals for single-colour movies and at 30-s intervals for two-colour and FRET movies. Automation was controlled by Andor IQ software. Two-colour movies of Crz1 and downstream targets were taken at 30 °C to ensure fast maturation of the Venus fluorophore.

Flow cytometry. Yeast cells were grown overnight in 96-well plates and diluted. Upon addition of calcium, cells were then regrown for 4–6 h and diluted fivefold. Gene expression was then measured using a Beckman Coulter Cell Lab Quanta SC MPL.

Image analysis. Fluorescence images were segmented using a Hough transformation algorithm in Matlab[48]. Localization score was determined by the difference between the mean intensity of the five brightest pixels in the cell and mean intensity of the rest of the pixels in the cell. Bursts were identified by thresholding traces at more than one standard deviation above background noise, estimated from the lowest 20% of values. Subsequent data analysis used the resulting traces with standard routines in Matlab and Mathematica.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 15 May; accepted 18 July 2008.

1. Cyert, M. S. Regulation of nuclear localization during signaling. *J. Biol. Chem.* **276,** 20805–20808 (2001).
2. Estruch, F. Stress-controlled transcription factors, stress-induced genes and stress tolerance in budding yeast. *FEMS Microbiol. Rev.* **24,** 469–486 (2000).

3. Kyriakis, J. M. The integration of signaling by multiprotein complexes containing Raf kinases. *Biochim. Biophys. Acta* **1773,** 1238–1247 (2007).
4. Stathopoulos-Gerontides, A., Guo, J. J. & Cyert, M. S. Yeast calcineurin regulates nuclear localization of the Crz1p transcription factor through dephosphorylation. *Genes Dev.* **13,** 798–803 (1999).
5. Yoshimoto, H. *et al.* Genome-wide analysis of gene expression regulated by the calcineurin/Crz1p signaling pathway in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **277,** 31079–31088 (2002).
6. Huh, W. K. *et al.* Global analysis of protein localization in budding yeast. *Nature* **425,** 686–691 (2003).
7. Mettetal, J. T. *et al.* The frequency dependence of osmo-adaptation in *Saccharomyces cerevisiae*. *Science* **319,** 482–484 (2008).
8. Hersen, P. *et al.* Signal processing by the HOG MAP kinase pathway. *Proc. Natl Acad. Sci. USA* **105,** 7165–7170 (2008).
9. Suel, G. M. *et al.* Tunability and noise dependence in differentiation dynamics. *Science* **315,** 1716–1719 (2007).
10. Ghaemmaghami, S. *et al.* Global analysis of protein expression in yeast. *Nature* **425,** 737–741 (2003).
11. Di Talia, S. *et al.* The effects of molecular noise and size control on variability in the budding yeast cell cycle. *Nature* **448,** 947–951 (2007).
12. Fewtrell, C. Ca^{2+} oscillations in non-excitable cells. *Annu. Rev. Physiol.* **55,** 427–454 (1993).
13. Wiesenberger, G. *et al.* Mg2+ deprivation elicits rapid Ca^{2+} uptake and activates Ca^{2+}/calcineurin signaling in *Saccharomyces cerevisiae*. *Eukaryot. Cell* **6,** 592–599 (2007).
14. Boustany, L. M. & Cyert, M. S. Calcineurin-dependent regulation of Crz1p nuclear export requires Msn5p and a conserved calcineurin docking site. *Genes Dev.* **16,** 608–619 (2002).
15. Roy, J. *et al.* A conserved docking site modulates substrate affinity for calcineurin, signaling output, and *in vivo* function. *Mol. Cell* **25,** 889–901 (2007).
16. Breuder, T. *et al.* Calcineurin is essential in cyclosporin A- and FK506-sensitive yeast strains. *Proc. Natl Acad. Sci. USA* **91,** 5372–5376 (1994).
17. Jacquet, M. *et al.* Oscillatory nucleocytoplasmic shuttling of the general stress response transcriptional activators Msn2 and Msn4 in *Saccharomyces cerevisiae*. *J. Cell Biol.* **161,** 497–505 (2003).
18. Garmendia-Torres, C., Goldbeter, A. & Jacquet, M. Nucleocytoplasmic oscillations of the yeast transcription factor Msn2: evidence for periodic PKA activation. *Curr. Biol.* **17,** 1044–1049 (2007).
19. Medvedik, O. *et al.* MSN2 and MSN4 Link calorie restriction and tor to sirtuin-mediated lifespan extension in *Saccharomyces cerevisiae*. *PLoS Biol.* **5,** e261 (2007).
20. Stathopoulos, A. M. & Cyert, M. S. Calcineurin acts through the CRZ1/TCN1-encoded transcription factor to regulate gene expression in yeast. *Genes Dev.* **11,** 3432–3444 (1997).
21. Golding, I. *et al.* Real-time kinetics of gene activity in individual bacteria. *Cell* **123,** 1025–1036 (2005).
22. Raj, A. *et al.* Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* **4,** e309 (2006).
23. Rodriguez, A. J. *et al.* Visualization of mRNA translation in living cells. *J. Cell Biol.* **175,** 67–76 (2006).
24. Elowitz, M. B. *et al.* Stochastic gene expression in a single cell. *Science* **297,** 1183–1186 (2002).
25. Bar-Even, A. *et al.* Noise in protein expression scales with natural protein abundance. *Nature Genet.* **38,** 636–643 (2006).
26. Cai, L., Friedman, N. & Xie, X. S. Stochastic protein expression in individual cells at the single molecule level. *Nature* **440,** 358–362 (2006).
27. Friedman, N., Cai, L. & Xie, X. S. Linking stochastic dynamics to population distribution: an analytical framework of gene expression. *Phys. Rev. Lett.* **97,** 168302-1–168302-4 (2006).
28. Kaern, M., Elston, T. C., Blake, W. J. & Collins, J. J. Stochasticity in gene expression: from theories to phenotypes. *Nature Rev. Genet.* **6,** 451–464 (2005).
29. Kaufmann, B. B. & van Oudenaarden, A. Stochastic gene expression: from single molecules to the proteome. *Curr. Opin. Genet. Dev.* **17,** 107–112 (2007).
30. Maheshri, N. & O'Shea, E. K. Living with noisy genes: how cells function reliably with inherent variability in gene expression. *Annu. Rev. Biophys. Biomol. Struct.* **36,** 413–434 (2007).
31. Newman, J. R. *et al.* Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* **441,** 840–846 (2006).
32. Ozbudak, E. M. *et al.* Regulation of noise in the expression of a single gene. *Nature Genet.* **31,** 69–73 (2002).
33. Sigal, A. *et al.* Variability and memory of protein levels in human cells. *Nature* **444,** 643–646 (2006).
34. Yu, J. *et al.* Probing gene expression in live cells, one protein molecule at a time. *Science* **311,** 1600–1603 (2006).
35. Rosenfeld, N. *et al.* Gene regulation at the single-cell level. *Science* **307,** 1962–1965 (2005).
36. Matheos, D. P. *et al.* Tcn1p/Crz1p, a calcineurin-dependent transcription factor that differentially regulates gene expression in *Saccharomyces cerevisiae*. *Genes Dev.* **11,** 3445–3458 (1997).
37. Armstrong, E. H. A method of reducing disturbances in radio signaling by a system of frequency modulation. *Proc. Inst. Radio Eng.* **24,** 689–740 (1936).
38. Song, G. B., Buck, N. V. & Agrawal, B. N. Spacecraft vibration reduction using pulse-width pulse-frequency modulated input shaper. *J. Guid. Control Dyn.* **22,** 433–440 (1999).
39. Adrian, E. D. & Zotterman, Y. The impulses produced by sensory nerve-endings: Part II. The response of a single end-organ. *J. Physiol. (Lond.)* **61,** 151–171 (1926).
40. Sarpeshkar, R. Analog versus digital: extrapolating from electronics to neurobiology. *Neural Comput.* **10,** 1601–1638 (1998).
41. Dolmetsch, R. E., Xu, K. & Lewis, R. S. Calcium oscillations increase the efficiency and specificity of gene expression. *Nature* **392,** 933–936 (1998).
42. Geva-Zatorsky, N. *et al.* Oscillations and variability in the p53 system. *Mol. Syst. Biol.* **2,** 2006.0033 (2006).
43. Nelson, D. E. *et al.* Oscillations in NF-κB signaling control the dynamics of gene expression. *Science* **306,** 704–708 (2004).
44. Friedman, N. *et al.* Precise temporal modulation in the response of the SOS DNA repair network in individual bacteria. *PLoS Biol.* **3,** e238 (2005).
45. Sheff, M. A. & Thorn, K. S. Optimized cassettes for fluorescent protein tagging in *Saccharomyces cerevisiae*. *Yeast* **21,** 661–670 (2004).
46. Gietz, R. D. & Woods, R. A. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol.* **350,** 87–96 (2002).
47. Sherman, F. Getting started with yeast. *Methods Enzymol.* **350,** 3–41 (2002).
48. Nachman, I., Regev, A. & Ramanathan, S. Dissecting timing variability in yeast meiosis. *Cell* **131,** 544–556 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Cyert for the CDRE and Crz1 mutant plasmids, K. Cunningham for the Crz1 overexpression plasmid pLE66, J. Stadler for the pGW845 FRET plasmid, S. Ramanathan for image analysis code, and K. Thorn, C.-L. Guo and L. LeBon for technical assistance. We thank U. Alon, M. Carlson, M. Cyert, H. Garcia, R. Kishony, G. Lahav, J.-G. Ojalvo, I. Riedel-Kruse, B. Shraiman, G. Süel, members of the laboratory, and especially N. Friedman for discussions. L.C. is supported by the Beckman Fellows Program at Caltech. This work was supported by National Institutes of Health grants R01GM079771 and P50 GM068763 for National Centers of Systems Biology, and the Packard Foundation.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.B.E. (melowitz@caltech.edu).

METHODS

Strain construction. All GFP strains were obtained from the GFP C-terminal protein fusion library (Invitrogen)[6]. All other strains were constructed similarly to the library: PCR was used to amplify cassettes with yeast codon-optimized fluorescent proteins tagged to an auxotrophic marker[35]. These amplified cassettes were made with primers containing 40 base-pair homology towards the carboxy terminus of the protein of interest. The cassettes were then transformed into the w303a laboratory strain of yeast using standard protocols[36]. Transformants were selected using standard yeast media[37]. Plasmid pGW845, a vector containing the FRET cameleon pair under the control of the ADH1 promoter, was provided by J. Stadler[13]; and p30, the Crz1^{high} mutant with altered calcineurin affinity, was provided by M. Cyert[15]. Both were transformed directly into a yeast strain containing Crz1–mCherry (Fig. 3b). A Crz1 overexpression plasmid, pLE66, provided by K. Cunningham[26], was transformed into each of 83 Crz1 target strains from the GFP library. A *Bam*HI fragment of yeast-codon-optimized Venus–YFP, cut from pKT103 (ref. 35), was ligated into pAMS342, pAMS363 and pAMS366, provided by M. Cyert, containing one, two and four CDREs respectively[20].

Media. Yeast were grown in synthetic complete or the appropriate drop-out media made using low-fluorescence yeast nitrogen, adapted from previous work[35]. This medium is yeast nitrogen base without riboflavin, folic acid or calcium chloride: 2 g l^{-1} $(NH_4)_2Cl_2$, 1 g l^{-1} KH_2PO_4, 0.5 g l^{-1} $MgCl_2$, 0.05 g l^{-1} NaCl, 0.5 mg l^{-1} H_3BO_4, 0.04 mg l^{-1} $CuSO_4$, 0.1 mg l^{-1} KI, 0.2 mg l^{-1} $FeCl_3$, 0.4 mg l^{-1} $MnCl_2$, 0.2 mg l^{-1} Na_2MoO_4, 0.4 mg l^{-1} $ZnSO_4$, 2 μg l^{-1} biotin, 0.4 mg l^{-1} calcium pantothenate, 2 mg l^{-1} inositol, 0.4 mg l^{-1} niacin, 0.2 mg l^{-1} PABA, 0.4 mg l^{-1} pyridoxine HCl, 0.4 mg l^{-1} thiamine, 20 g l^{-1} dextrose.

Time-lapse microscopy. Cells were attached to glass bottom dishes (Willco) functionalized with concanavalin-A (Sigma). Fluorescence images were taken at room temperature on an Olympus IX81 with the ZDC autofocus option and an Andor Ikon (DU-934) camera. $CaCl_2$ was added during movies to nominal concentrations using a 2× stock solution. Images were acquired at 10-s intervals for single-colour movies and at 30-s intervals for two-colour and FRET movies. Automation was controlled by Andor IQ software. Two-colour movies of Crz1 and downstream target genes were taken at 30 °C to ensure fast maturation of the Venus fluorophore.

Flow cytometry. Yeast cells were grown overnight in 96-well plates and diluted. $CaCl_2$ was added to nominal concentrations with a 2× stock solution. Cells were regrown in the calcium for 4–6 h, diluted fivefold, and gene expression was measured using a Beckman Coulter Cell Lab Quanta SC MPL. To identify cells, all particles were triggered and gated during acquisition. As shown in Supplementary Fig. 13, particles were triggered if their Coulter volume measurements were within a specified range. Either 5,000 particles or a 30-μl sample volume were acquired. These particles were then excited by a 22-mW 488-nm laser, and side-scatter and fluorescence were measured. A polygon gate was drawn on the Coulter volume/side-scatter two-parameter plot to identify cells. If more than 500 cells were found, mean fluorescence was calculated. Quanta photomultiplier tube settings were 3.5 for side-scatter, 8.00 for GFP measurements and from 4 to 6 for synthetic promoter–YFP measurements.

Image analysis. Fluorescence cell images were segmented using a Hough transformation algorithm in Matlab, provided by S. Ramanathan[38]. Localization score was determined by the difference between the mean intensity of the five brightest pixels in the cell and mean intensity of the rest of the pixels in the cell. Bursts were identified by thresholding traces at more than one standard deviation above background noise, estimated from the lowest 20% of values. Subsequent data analysis used the resulting traces with standard routines in Matlab and Mathematica.

ARTICLES

Trans-splicing in *C. elegans* generates the negative RNAi regulator ERI-6/7

Sylvia E. J. Fischer[1]*, Maurice D. Butler[1]*, Qi Pan[1]† & Gary Ruvkun[1]

Mutations that enhance the response to double-stranded RNA (dsRNA) have revealed components of the RNA interference (RNAi) pathway or related small RNA pathways. To explore these small RNA pathways, we screened for *Caenorhabditis elegans* mutants displaying an enhanced response to exogenous dsRNAs. Here we describe the isolation of mutations in two adjacent, divergently transcribed open reading frames (*eri-6* and *eri-7*) that fail to complement. *eri-6* and *eri-7* produce separate pre-messenger RNAs (pre-mRNAs) that are *trans*-spliced to form a functional mRNA, *eri-6/7*. *Trans*-splicing of *eri-6/7* is mediated by a direct repeat that flanks the *eri-6* gene. Adenosine to inosine editing within untranslated regions of *eri-6* and *eri-7* pre-mRNAs reveals a double-stranded pre-mRNA intermediate, forming in the nucleus before splicing occurs. The ERI-6/7 protein is a superfamily I helicase that both negatively regulates the exogenous RNAi pathway and functions in an endogenous RNAi pathway.

RNAi pathways act throughout phylogeny as both an experimental gene-silencing tool and a regulator of endogenous gene expression[1,2]. The endonuclease Dicer[3] and various Argonaute proteins act in specialized roles in most silencing pathways[4]. *C. elegans* Dicer interacts with negative regulators of exogenous RNAi[5], such as the RNA-dependent RNA polymerase (RdRP) RRF-3 (ref. 6) and the exonuclease ERI-1 (ref. 7). These proteins are required for the production or stability of a subset of endogenous short interfering RNAs (siRNAs), suggesting a competition with the exogenous RNAi pathway for shared, rate-limiting factors[5,8].

RdRPs amplify siRNAs on mRNA templates in nematodes[9–11], fungi and plants[1]. The feedforward nature of RNAi and the still unexplained resistance of neurons to RNAi[12–14] suggest that there are undiscovered negative regulators of RNAi. To identify such regulators, we conducted a genetic screen for mutations that confer an enhanced response to dsRNA. Here we describe the identification and characterization of adjacent genes, *eri-6* and *eri-7*, which are assembled by a dsRNA-mediated *trans*-splicing mechanism to regulate RNAi in a negative manner. The dsRNA-dependent production of an RNAi factor suggests that there may be an autoregulatory feedback mechanism in RNAi.

Characterization of *eri* mutants

To identify negative regulators of *C. elegans* RNAi, we carried out a genetic screen for mutants having an enhanced RNAi (Eri) phenotype. These mutants, unlike wild type, exhibited an enhanced response to green fluorescent protein (*gfp*) RNAi, downregulating expression of a neuronally expressed *gfp* transgene (*unc-47::gfp*)[15], and an enhanced response to dsRNA targeting several non-neuronally expressed endogenous genes, including the collagen *dpy-13* (ref. 16), the cadherin *hmr-1* (ref. 17) and the transcription factor *lin-1* (ref. 18). Thirteen of the 44 *eri*(−) mutants were also thermosensitive sterile and showed X-chromosome non-disjunction, phenotypes that define a class of *eri* genes that encode the DCR-1-interacting proteins RRF-3, ERI-1, ERI-3 and ERI-5 (refs 5–7). Among the mutations not in this class, *mg379* and the independently isolated spontaneous *eri* allele *mg411* enhance RNAi and silence a transgene (*mgIs30*) that confers a rolling locomotion (Rol) phenotype (Supplementary Fig. 2 and Supplementary Tables 1 and 2). Transgene silencing and exogenous RNAi are mediated by several common factors[19], and all known *eri*(−) mutants silence transgenes[6,7]. *mg379* and *mg411*, as well as other allele combinations, fail to complement (Supplementary Table 1 and Supplementary Information). *mg411* and *mg379* were mapped to a 1.7-megabase interval. The 330 genes in the interval were assayed for a loss-of-function phenotype of transgene silencing by using dsRNAs that target each gene[20]. Two dsRNA clones induced silencing of the *mgIs30* transgene, phenocopying the *mg411* and *mg379* mutations. These dsRNAs target two adjacent, antiparallel open reading frames: C41D11.7, termed *eri-7*, and C41D11.1, termed *eri-6* (Fig. 1a). These RNAi clones are specific for each gene and do not target the adjacent gene or any other gene in the genome. DNA sequencing identified *mg411* as a missense mutation in C41D11.7 and *mg379* as a 5′ splice site mutation in C41D11.1. DNA sequencing of mutants isolated in the genetic screen revealed five more mutations in C41D11.7: *mg369*, *mg380*, *mg390*, *mg440* and *mg442*. An independently isolated deletion allele (*mg441*) removes *eri-6* (Supplementary Fig. 3). The transgene-silencing phenotype of mutants with either *eri-7*(*mg411*) or *eri-6*(*mg379*) alleles was rescued by transformation of a 9.1-kilobase (kb) fragment that contains only C41D11.1 and C41D11.7 (Fig. 2a, b). The hypersensitivity to RNAi by feeding conferred by the *mg411* allele was also rescued by the 9.1-kb fragment (data not shown).

Anomalous *eri-6/7* gene structure

Although the lack of complementation suggests that *eri-6*(*mg379*) and *eri-7*(*mg411*) affect a single gene, the mutations were found in ostensibly separate genes on antiparallel strands[21]. However, the syntenic region in the nematode species *C. briggsae* surprisingly revealed that *eri-6* and *eri-7* constitute a single, contiguous gene in this species (Fig. 1a). The 5′ coding region of the *C. briggsae* gene CBG03999 is homologous to *eri-6*, whereas the 3′ coding region is homologous to *eri-7*. We verified the *C. elegans* genome structure at the *eri-6* and *eri-7* loci by Southern blotting

[1]Department of Molecular Biology, Massachusetts General Hospital, and Department of Genetics, Harvard Medical School, Boston, Massachusetts 02114, USA. †Present address: ZymoGenetics, 1201 Eastlake Avenue East, Seattle, Washington 98102, USA.
*These authors contributed equally to this work.

(Supplementary Fig. 4), and by PCR and sequencing analyses (data not shown). We investigated closer relatives of the *C. elegans* wild-type N2 strain—27 other natural isolates of *C. elegans*. Four of these, including an isolate from Hawaii, *C. elegans* CB4856 (ref. 22), show a fused gene structure similar to that of *C. briggsae* (Fig. 1b).

eri-6 in the *C. elegans* strain N2 is flanked by a nearly perfect, approximate 930-base-pair (bp) direct repeat (Fig. 1b). Within this direct repeat is a 25-bp inverted repeat (Fig. 1b and Supplementary Fig. 5). The inverted repeat may have mediated the rearrangement responsible for the difference in *eri-6* orientation relative to *eri-7* in *C. elegans* N2 compared with *C. elegans* CB4856 (Fig. 1b). This hypothesis is supported by the spontaneous allele *mg441* (Supplementary Fig. 3 and Supplementary Discussion), which deletes *eri-6* in a fashion consistent with a rearrangement having occurred through the approximate 930-bp direct repeats that flank *eri-6*.

eri-6 and *eri-7* are *trans*-spliced

Sequencing of the expressed sequence tag (EST) yk224h9, and reverse transcription-PCR (RT–PCR), showed that the divergently transcribed RNAs for *eri-6* and *eri-7* are assembled into one mRNA in *C. elegans* strain N2 (Fig. 1a), encoding a protein orthologous to the *C. briggsae eri-6/7* gene product. These proteins are members of the superfamily I DNA and RNA helicases (Supplementary Figs 6 and 7). Several *eri-6/7* mutations encode amino-acid substitutions in conserved motifs in the ATPase and helicase domains (Supplementary Fig. 7). A transgene containing the *eri-6* promoter, the full-length *eri-6/7* complementary DNA (cDNA), and the *eri-7* 3′ untranslated region (UTR) rescues both the *eri-6*(*mg379*) and the *eri-7*(*mg411*) phenotypes (Supplementary Table 3).

In addition to the helicase mRNA, three other *eri-6* splice variants were identified by EST analysis and 3′ rapid amplification of cDNA ends (RACE), in which *eri-6* splices to different downstream exons (Fig. 1a), encoding proteins with no homology to known proteins. Neither the *eri-6* mRNA nor the *eri-7* mRNA could be detected as the separate mRNAs predicted from the gene structure.

We considered the following two mechanisms of chimaeric RNA formation: (1) *trans*-splicing of independently transcribed pre-mRNAs; and (2) a low level of DNA inversion in particular cells that would reposition the *eri-6* and *eri-7* genes *in cis*. The genomic rearrangement model was alluring because of the observation that such inversions at the *eri-6/7* locus are apparent between and within related *Caenorhabditis* species. However, the rearrangement model is neither supported by Southern blotting (Supplementary Fig. 4) nor extensive PCR analyses designed to detect a *cis* orientation of *eri-6* and *eri-7* as a minority component of the genomic DNA of *C. elegans* N2 (Supplementary Discussion and Supplementary Fig. 8).

The generation of the *eri-6/7* mRNA by *trans*-splicing is, however, supported by multiple experimental data. The *trans*-splicing juncture

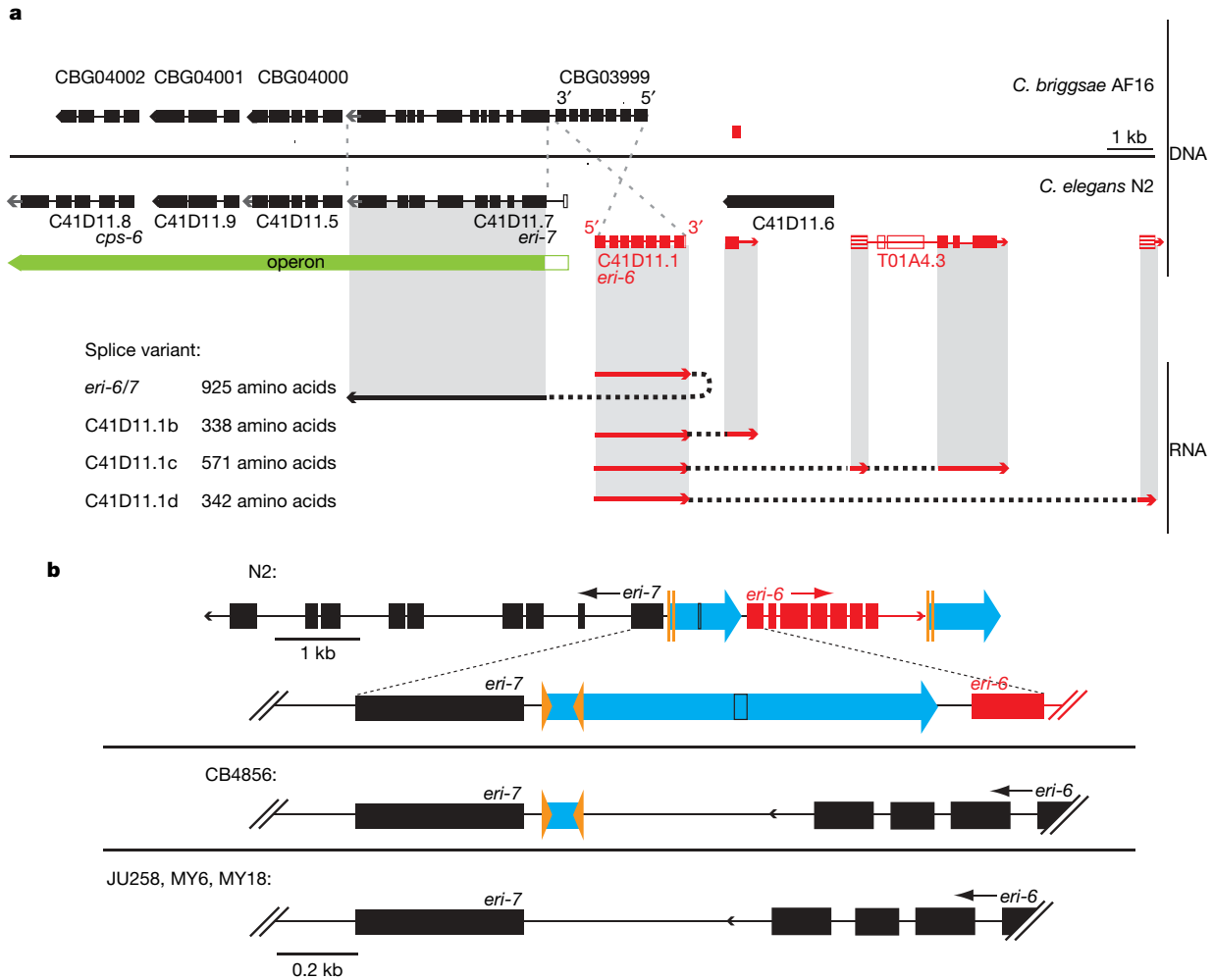


Figure 1 | The *eri-6/7* locus in various strains and its gene products in *C. elegans* N2. **a**, Structure of *eri-6/7* in *C. elegans* N2 and the syntenic region in *C. briggsae* AF16, based on RT–PCR, 3′-RACE and WormBase[21]. The *eri-7* operon is conserved. *C. briggsae eri-6* is encoded on the opposite strand from *C. elegans* and is one gene with *C. briggsae eri-7* (CBG03999). Exons are black or red rectangles, depending on the strand on which they are encoded. Open rectangles: predicted exons[21], but unconfirmed. Striped exons: discovered experimentally. *eri-6* splices to *eri-7* 24 nucleotides upstream of the predicted[21] *eri-6* stop codon. **b**, Relative orientation of *eri-6* and *eri-7* in four *C. elegans* isolates. Blue arrows: the approximate 930-bp direct repeat flanking *eri-6*. Orange arrowheads: 25-bp inverted repeats.

of the fused *eri-6/7* mRNA uses canonical *cis* splice sites (Supplementary Fig. 9). Other *eri-6* splice variants were detected 3′ of *eri-6* and all include exons linked within 12 kb, *in cis* (Fig. 1a). Bioinformatic searches (Supplementary Discussion) of EST databases and 5′-RACE experiments did not detect distant *trans*-splicing of *eri-6* to other mRNAs. Therefore, it is probable that for *trans*-splicing to occur, the *eri-6* and *eri-7* genes need to be in close proximity. Genes can also be juxtaposed through homologous chromosome pairing during meiosis[23]. Interchromosomal *trans*-splicing of *eri-6* and *eri-7* is unlikely to efficiently occur given (1) the lack of complementation observed in *eri-7*(+)*eri-6*(*mg379*)/*eri-7*(*mg369*)*eri-6*(+) animals, and (2) sequence analysis of cDNAs cloned from these *trans*-heterozygotes (Supplementary Discussion).

A prerequisite for co-transcriptional *trans*-splicing is expression of both pre-mRNAs in the same cells at the same time. *eri-7* promoter::*gfp* and *eri-6* promoter::*rfp* transgenes are expressed in overlapping patterns, with co-expression in hypodermal cells and two pairs of sensory head neurons (ASK and ASI; Fig. 2c). The *eri-7* promoter also expresses in the somatic gonad.

The direct repeat flanking *eri-6* (Fig. 1b) could constitute the 5′-UTR of the *eri-7* pre-mRNA and, as a reverse complement copy, the *eri-6* pre-mRNA 3′- or 5′-UTR. These complementary sequences could bring the pre-mRNAs into physical contact to facilitate *trans*-splicing. To determine the pre-mRNA sequences of *eri-6* and *eri-7*, we stabilized them by attenuating splicing through RNAi inactivation of *rnp-5*, which encodes an orthologue of human RNPS1 (ref. 21), a member of the exon-junction complex. 5′-RACE analysis using intron-specific primers identified an *eri-7* pre-mRNA 5′ end extending 725 bp into the *eri-6/7* intergenic region (685 bp into the first direct repeat). An *eri-6* transcription start site was found at up to 100 bp 5′ of the *eri-6* start codon (Fig. 2d). We were unable to amplify the *eri-6* 3′-UTR, possibly because of a repetitive, low complexity

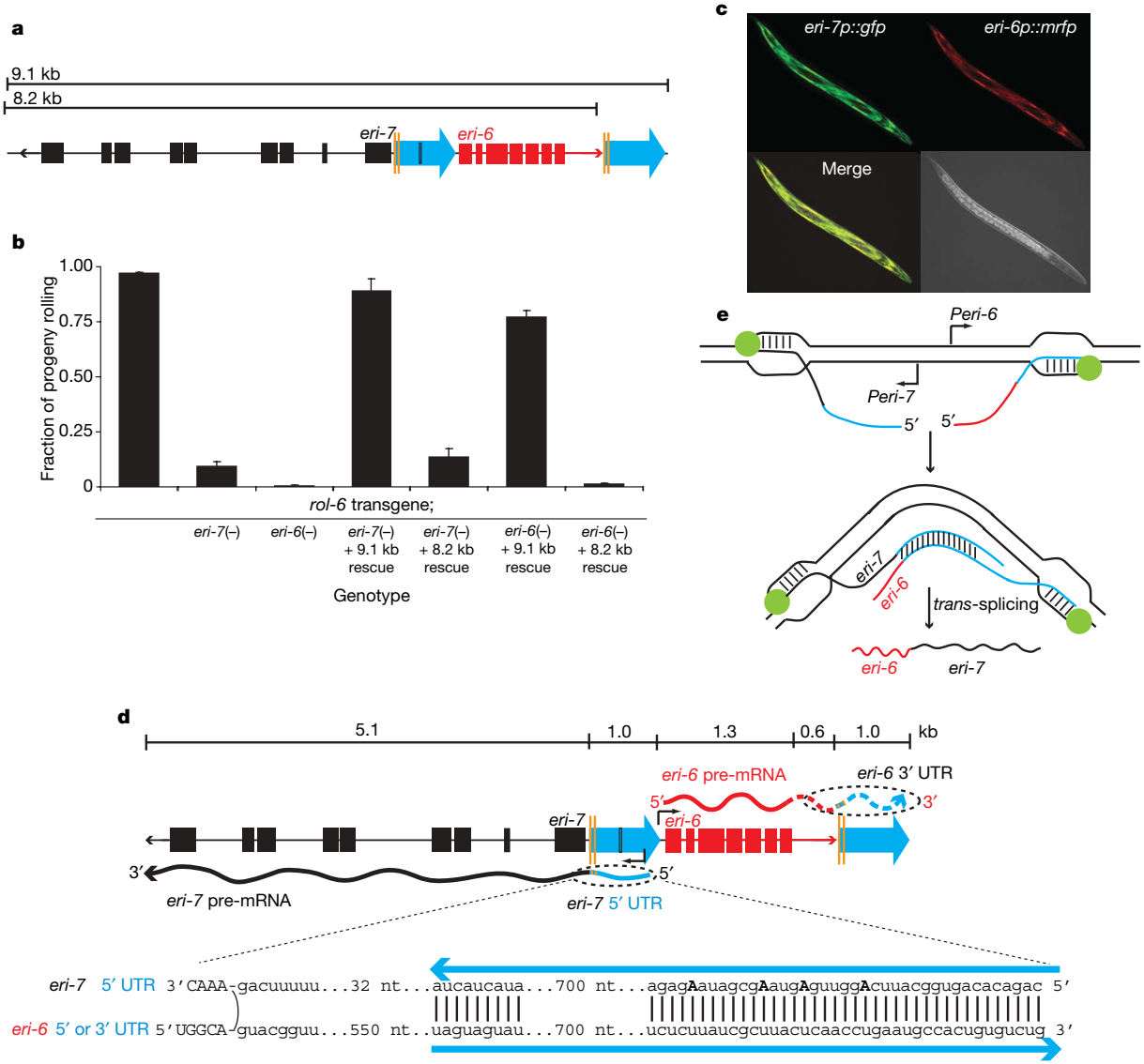


Figure 2 | The *eri-6/7* mRNA is formed by local *trans*-splicing. **a**, Schematic of *eri-6/7* locus used for rescue: 9.1-kb fragment; and 8.2-kb fragment missing the second repeat. **b**, Rescue of *eri-6/7*(*mg411*)-induced silencing of the *mgIs30* transgene-conferred rolling phenotype. Number of broods scored = 28, 28, 18, 19, 38, 29 and 18, respectively. Multiple transgenic lines assayed. Error bars: s.e.m. **c**, The *eri-6* and *eri-7* promoters express in overlapping tissues (hypodermis and head neurons). *eri-6* and *eri-7* promoters were fused to *rfp* and *gfp*, respectively. **d**, 5′-RACE experiments show that *eri-6* and *eri-7* are expressed as separate pre-mRNAs with sufficient nucleotide homology to base-pair, thus facilitating *trans*-splicing. *eri-7* pre-mRNA starts 775 nucleotides (nt) upstream of exon 1, whereas *eri-6* pre-mRNA starts 100 nt upstream of *eri-6* exon 1 (black arrows). Edited nucleotides (A to G indicative of A to I) in *eri-7* pre-mRNA are in bold capital letters. Blue arrows: direct repeats. Capital letters: exon sequence. Curved line: *trans*-splicing juncture. **e**, Model of *eri-6/7 trans*-splicing. *eri-6* and *eri-7* pre-mRNAs are locally co-transcribed in the same cells. The pre-mRNAs may form a dsRNA intermediate that facilitates co-transcriptional *trans*-splicing.

sequence immediately 3′ of the splice site (5′ of the second direct repeat). The splicing of *eri-6* to independent exons located 1–11 kb 3′ of the second direct repeat strongly supports an *eri-6/7 trans*-splicing model whereby the *eri-6* nascent transcript bears this repeat.

The significance of the second 930-bp direct repeat was tested in rescue experiments using an *eri-6/7* (9.1-kb) genomic fragment that contained this repeat, compared with an (8.2-kb) fragment that did not (Fig. 2a, b). The 9.1-kb fragment completely rescued the transgene-silencing phenotypes of both the *eri-6*(*mg379*) and *eri-7*(*mg411*) alleles, whereas the 8.2-kb fragment only weakly rescued the *eri-7* mutation. This nominal *eri-6/7* gene activity could be the result of overexpression of *eri-6* and *eri-7* from the transgenic array bypassing the need for their base pairing, or of a low level of *eri-6* transcription from an upstream start site, to include repeat sequence in the 5′-UTR that could mediate base-pairing.

Analysis of cDNA sequences derived from the *eri-7* pre-mRNA stabilized in *rnp-5* RNAi-treated animals revealed adenosine to guanosine transitions at four positions located within the direct repeat (Fig. 2d). These transitions are indicative of adenosine to inosine editing of the *eri-7* 5′-UTR by an adenosine deaminase (ADAR)[24], supporting the model that the direct repeat element flanking *eri-6* is bidirectionally transcribed to form an RNA duplex intermediate before the two mRNAs are *trans*-spliced (Fig. 2e). Although ADARs can also edit secondary structures within single RNA strands, no such structure is predicted for the *eri-7* 5′-UTR.

To show that the *trans*-spliced mRNA is translated into a chimaeric protein, we constructed a *trans*-splicing reporter in which all but the initial four nucleotides of *eri-7* sequence are replaced by a *gfp* gene lacking a start codon (Supplementary Fig. 10). Sequencing of RT–PCR products (Fig. 3a) from GFP-expressing worms (Fig. 3b) showed that *trans*-splicing occurred from the transgene to yield an mRNA that consists of *eri-6*, the first two codons of *eri-7*, and *gfp* sequence. Western blotting showed that a protein of the size expected for the ERI-6::GFP fusion can be detected using a GFP antibody, indicating that the *eri-6* and *eri-7 trans*-spliced mRNA is probably translated into a chimaeric ERI-6/7 helicase protein (Fig. 3c).

This reporter for *eri-6/7 trans*-splicing is expressed in the amphid neuron ASK (Fig. 3b) and in the somatic gonad, but not in the hypodermis where the promoter fusions are strongly expressed. A second transgene, bearing the complete *eri-6/7* genomic region, fuses mCherry[25] to the carboxy terminus of ERI-7 (Fig. 3d–f and Supplementary Fig. 11), and shows expression in ASK and the somatic gonad. One explanation for the lack of hypodermal expression from these *trans*-splicing-dependent fusion genes is that *trans*-splicing is inefficient or inhibited in the hypodermis. Alternatively, transcription or translation in the hypodermis does not occur because of sequence in the *eri-6/7* gene or mRNA that is not present in the promoter fusion constructs. Interestingly, the *eri-6/7 trans*-splicing reporter is more highly and broadly expressed when RNAi is attenuated by inactivation of the Argonaute gene *rde-1* (Supplementary Fig. 12), suggesting that the *eri-6/7* is a target of RNAi. A caveat to this result is that transgenes are generally expressed at higher levels when RNAi is defective. Quantitative RT–PCR revealed that *eri-6/7* mRNA levels are not increased in animals lacking RDE-1 (data not shown). One explanation for these data is that the dsRNA element of *eri-6/7* acts tissue-specifically as a sensor, and the changes in expression are too subtle to detect by quantitative RT–PCR on whole worms.

C. elegans has two ADAR genes: *adr-1* and *adr-2*. The *adr-1*(−); *adr-2*(−) double mutant enhances transgene silencing, suggesting that editing of dsRNAs normally inhibits their sensitivity to RNAi[26]. Perhaps in the absence of ADARs, the unedited *eri-6/7* dsRNA intermediate is a better target for the RNAi pathway, thus decreasing *eri-6/7* gene activity. However, the level of *trans*-spliced *eri-6/7* mRNA is unchanged in the *adr-1*(−); *adr-2*(−) double mutant as analysed by quantitative RT–PCR and by GFP fluorescence of the *trans*-splicing reporter after *adr* RNAi (data not shown), suggesting that editing of the double-stranded intermediate does not affect *eri-6/7 trans*-splicing.

Role of ERI-6/7 in RNAi

eri-6/7 alleles define a new class of Eri mutants. They lack the pleiotropic phenotypes (for example, sterility) typical of when function of the Dicer-interactors ERI-1 to ERI-5 is lost. In addition, *eri-6/7*(−) mutants do not show the characteristic synthetic multivulva phenotype that is associated with enhanced RNAi mutants of the retinoblastoma pathway[27]. We analysed the role of *eri-6/7* in small RNA pathways. Like *eri-1* alleles[7], *eri-7*(*mg369*) confers an enhanced response (over wild type) to injected double-stranded siRNAs (*unc-22*) (Supplementary Table 1); however, we did not observe a significantly increased level of siRNAs corresponding to a target gene after exposure to dsRNAs as analysed by northern blotting (data not shown). These data suggest that *eri-6/7* acts downstream of processing of the exogenous trigger-dsRNA into siRNAs. Genetic epistasis analyses (Supplementary Table 1 and data not shown) with mutants defective in RNAi (*rde-1* (ref. 19), *rde-4* (ref. 28), *rrf-1* (ref. 11) and *mut-7* (ref. 29)) show that, with the exception of the somatic RdRP gene *rrf-1*,

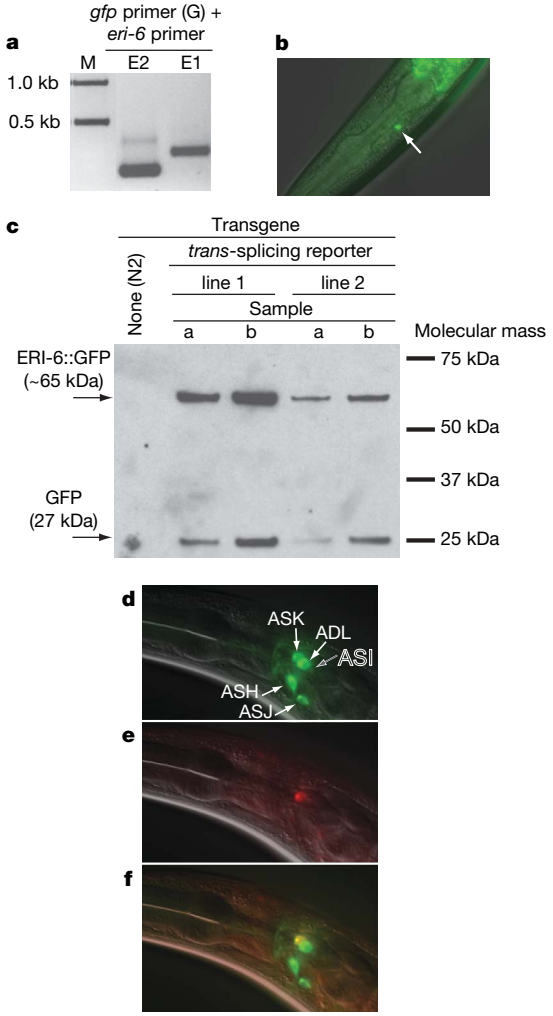


Figure 3 | Chimaeric ERI-6/7 protein is expressed through *trans*-splicing. **a**, RT–PCR on *gfp*-expressing worms using primers in *gfp* (G) and *eri-6* (E1 and E2). Sequencing of the RT–PCR products confirmed *trans*-splicing. **b**, Expression of the *trans*-splicing reporter in ASK. **c**, Western blot on animals expressing the *trans*-splicing reporter. A GFP antibody was used. The slower-migrating product (of the expected weight of ERI-6::GFP) is presumably ERI-6::GFP. The faster-migrating product may be GFP expressed using an alternative start codon. **d**, DiO fills five pairs of sensory neurons. **e**, ERI-6/7::mCherry is expressed in one amphid neuron. **f**, Merged image identifies ASK as the ERI-6/7-expressing neuron.

normal function of these genes is required for enhanced and normal RNAi sensitivity. RRF-1 functions in the *de novo* production of secondary siRNAs after the initial cleavage of the exogenous dsRNA into primary siRNAs by Dicer, RDE-1 and RDE-4 (ref. 11). Unlike *rrf-1*(−) mutants, *rrf-1*(−) *eri-7*(−) double mutants are partly proficient in somatic RNAi targeting *hmr-1* and *unc-22*, suggesting that *eri-6/7* negatively acts both in *rrf-1*-independent RNAi (through primary siRNAs) and in *rrf-1*-dependent RNAi. These data are consistent with a role of ERI-6/7 downstream of, and dependent on, the siRNA production mediated by RDE-1 and RDE-4. Null alleles of *eri-1* or *rrf-3*, when combined with *eri-6/7* null alleles, did not enhance the RNAi sensitivity compared with the single mutants, suggesting that *eri-6/7* may act in the same pathway as *eri-1* and *rrf-3* to enhance RNAi (data not shown).

Silencing in response to exogenous dsRNA is increased when certain endogenous RNAi pathways, such as the pathway dependent on *eri-1*, are defective[5]. We investigated a role for ERI-6/7 in endogenous RNAi by analysing endogenous siRNA (endo-siRNA) abundance in *eri-6/7* mutant worms by northern blotting (Fig. 4). *eri-6/7* is required for endo-siRNAs derived from the gene K02E2.6 but not for a Dicer-independent endo-siRNA matching the gene T01A4.3 (Fig. 4), which is thought to be germline-expressed[5]. ERI-6/7 protein shares these attributes with the DCR-1-interacting proteins ERI-1, RRF-3 and RDE-4 (refs 5 and 8; Fig. 4), as well as with the Argonaute ERGO-1 (ref. 30).

Conclusions

eri-6/7 encodes an RNAi factor that is assembled in *C. elegans* N2 by *trans*-splicing through a dsRNA intermediate. The *trans*-splicing of *eri-6* and *eri-7* is only observed when the two genes are on the same chromosome, suggesting that it occurs on nascent transcripts. The *eri-6/7* double-stranded, pre-mRNA intermediate is edited by ADARs when splicing is inhibited, and may be itself regulated by RNAi.

ERI-6/7 protein is a member of the superfamily I helicases, like the RNAi factors Mov10 in mammals[31], SDE3 in *Arabidopsis*[32] and Armitage in *Drosophila*[33]. Superfamily I helicases act on either RNA or DNA. Because ERI-6/7 is required for endo-siRNA production or stability, it could function as an RNA helicase in siRNA generation. Our expression data showing exclusive or predominant cytoplasmic expression of a full-length ERI-6/7 protein fused to a fluorescent marker (Supplementary Fig. 11) suggest that it acts as an RNA helicase.

Loss of *eri-6/7* causes enhanced exogenous RNAi. Lee *et al.* (2006) and Duchaine *et al.* (2006) have proposed that other ERI proteins inhibit exogenous RNAi indirectly by competing for factors common to both the endogenous and exogenous RNAi pathways, suggesting that the normal role of ERI-6/7 is either to promote endogenous small RNA pathways or to attenuate exogenous RNAi directly. The loss of some endogenous siRNAs in *eri-6/7* mutants is consistent with the former function.

This is the first example of exon-to-exon *trans*-splicing in *C. elegans*, and only the third finding of a locus that requires such *trans*-splicing for the production of a functional protein. The other two examples were reported in *Drosophila*: *mod*(*mdg4*)[34,35] and *lola*[36]. Each of these genes encodes BTB–zinc-finger transcription factors that achieve extensive protein diversity through alternative *trans*-splicing. The requirement of proximity between the *trans*-spliced exons in these cases is not clear[35,36].

The polymorphic variation of the *eri-6/7* locus between related *Caenorhabditis* species, and especially between different isolates of the *C. elegans* species, is remarkable. The inserted repetitive sequence has probably segmented the *eri-6/7* gene into two in *C. elegans* N2 through recombinational events. The ability of the repeat to mediate homology-based *trans*-splicing, and its bidirectional promoter activity (allowing independent *eri-6* and *eri-7* expression in N2), have provided a compensatory means of bringing the two fragments together again to encode a composite *eri-6/7* mRNA. Beyond the structure of the *eri-6/7* locus, the ERI-6/7 proteins seem to have also diverged substantially between *C. elegans* and *C. briggsae*. Whereas the mean amino-acid identity between orthologues is 75% (ref. 37), ERI-6/7 shows only 44% identity to its fused orthologue in the related nematode *C. briggsae*. This is consistent with the rapid evolution that is seen in antiviral RNAi genes[38], and with the finding that other *C. elegans eri* genes are anti-viral[39,40].

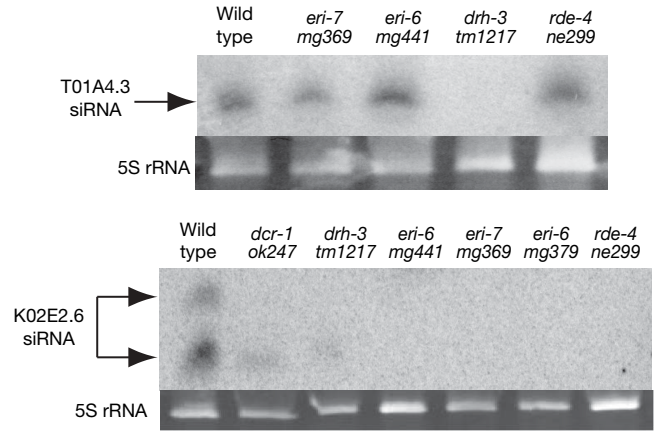


Figure 4 | ERI-6/7 is required for endogenous RNAi. Shown are northern blots probed for two different endogenous siRNAs, targeting the genes T01A4.3 and K02E2.6. Mutants of additional genes known to be required for endogenous RNAi (*dcr-1*, *drh-3* and *rde-4*) are shown as controls. 5S rRNA is shown as a loading control.

METHODS SUMMARY

C. elegans was cultured using standard techniques[41] and fed on *Escherichia coli* OP50 or on *E. coli* HT115 harbouring a dsRNA expressing plasmid (RNAi by feeding).

C. elegans animals were mutagenized using ethylmethanesulphonate. Mutants showing an enhanced RNAi phenotype were selected after *gfp* RNAi in the F_2 and F_3 progeny of mutagenized animals carrying a transgene expressing GFP in a subset of neurons. The enhanced RNAi phenotype was further analysed using RNAi to the endogenous genes *hmr-1*, *lin-1* and *dpy-13*. In addition, several mutants were assayed for their ability to silence the transgene *mgIs30* (*lim-6::gfp*, *col-10::lacZ::lin-14*, *rol-6*(*su1006*)*::lin-14* 3′-UTR), scoring the Rol phenotype conferred by this transgene. *unc-22* siRNA injections were done as described previously[7,42].

Mutations were single-nucleotide-polymorphism-mapped using the Eri and transgene silencing phenotypes. Subsequently, all genes in the mapping interval were inactivated by RNAi and tested for transgene silencing. Inactivation of two genes silenced the transgene *mgIs30*; these genes were sequenced in the mutants.

Transgenic worms for rescue experiments, expression analysis as well as animals carrying the reporter for *trans*-splicing were made by microinjection of PCR fragments, cosmids or plasmids. Promoter and full-length fusions to fluorescent proteins were made by splicing by overlapping extension PCR of *gfp*, *mrfp* or *mCherry* to the promoter or in frame with the open reading frame. Western blotting of the *eri-6::gfp* reporter was done using an anti-GFP antibody.

The gene structure or *eri-6/7* in wild-type *C. elegans* N2, other *C. elegans* isolates and *C. briggsae* was confirmed by PCR, sequencing and/or Southern blotting.

eri-6/7 transcripts were analysed by EST sequencing analysis, RT–PCR, 5′- and 3′-RACE. cDNA was cloned and sequenced. Pre-mRNAs and transcriptional start sites were analysed in animals that were deficient in splicing by RNAi of *rnp-5*.

Endogenous siRNA abundance in mutant animals was assayed by northern blotting[42] using StarFire probes[5].

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 9 May; accepted 18 July 2008.
Published online 10 September 2008.

1. Mello, C. C. & Conte, D. Jr. Revealing the world of RNA interference. *Nature* 431, 338–342 (2004).

2. Zamore, P. D. & Haley, B. Ribo-gnome: the big world of small RNAs. *Science* **309,** 1519–1524 (2005).
3. Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409,** 363–366 (2001).
4. Parker, J. S. & Barford, D. Argonaute: a scaffold for the function of short regulatory RNAs. *Trends Biochem. Sci.* **31,** 622–630 (2006).
5. Duchaine, T. F. *et al.* Functional proteomics reveals the biochemical niche of *C. elegans* DCR-1 in multiple small-RNA-mediated pathways. *Cell* **124,** 343–354 (2006).
6. Simmer, F. *et al.* Loss of the putative RNA-directed RNA polymerase RRF-3 makes *C. elegans* hypersensitive to RNAi. *Curr. Biol.* **12,** 1317–1319 (2002).
7. Kennedy, S., Wang, D. & Ruvkun, G. A conserved siRNA-degrading RNase negatively regulates RNA interference in *C. elegans*. *Nature* **427,** 645–649 (2004).
8. Lee, R. C., Hammell, C. M. & Ambros, V. Interacting endogenous and exogenous RNAi pathways in *Caenorhabditis elegans*. *RNA* **12,** 589–597 (2006).
9. Alder, M. N., Dames, S., Gaudet, J. & Mango, S. E. Gene silencing in *Caenorhabditis elegans* by transitive RNA interference. *RNA* **9,** 25–32 (2003).
10. Pak, J. & Fire, A. Distinct populations of primary and secondary effectors during RNAi in *C. elegans*. *Science* **315,** 241–244 (2007).
11. Sijen, T. *et al.* On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell* **107,** 465–476 (2001).
12. Kamath, R. S. *et al.* Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* **421,** 231–237 (2003).
13. Tavernarakis, N., Wang, S. L., Dorovkov, M., Ryazanov, A. & Driscoll, M. Heritable and inducible genetic interference by double-stranded RNA encoded by transgenes. *Nature Genet.* **24,** 180–183 (2000).
14. Timmons, L., Court, D. L. & Fire, A. Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*. *Gene* **263,** 103–112 (2001).
15. McIntire, S. L., Reimer, R. J., Schuske, K., Edwards, R. H. & Jorgensen, E. M. Identification and characterization of the vesicular GABA transporter. *Nature* **389,** 870–876 (1997).
16. von Mende, N., Bird, D. M., Albert, P. S. & Riddle, D. L. *dpy-13*: a nematode collagen gene that affects body shape. *Cell* **55,** 567–576 (1988).
17. Costa, M. *et al.* A putative catenin–cadherin system mediates morphogenesis of the *Caenorhabditis elegans* embryo. *J. Cell Biol.* **141,** 297–308 (1998).
18. Beitel, G. J., Tuck, S., Greenwald, I. & Horvitz, H. R. The *Caenorhabditis elegans* gene *lin-1* encodes an ETS-domain protein and defines a branch of the vulval induction pathway. *Genes Dev.* **9,** 3149–3162 (1995).
19. Tabara, H. *et al.* The *rde-1* gene, RNA interference, and transposon silencing in *C. elegans*. *Cell* **99,** 123–132 (1999).
20. Fraser, A. G. *et al.* Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference. *Nature* **408,** 325–330 (2000).
21. WormBase Release WS170. ⟨http://ws170.wormbase.org/⟩ (2007).
22. Wicks, S. R., Yeh, R. T., Gish, W. R., Waterston, R. H. & Plasterk, R. H. Rapid gene mapping in *Caenorhabditis elegans* using a high density polymorphism map. *Nature Genet.* **28,** 160–164 (2001).
23. Garcia-Muse, T. & Boulton, S. J. Meiotic recombination in *Caenorhabditis elegans*. *Chromosome Res.* **15,** 607–621 (2007).
24. Bass, B. L. RNA editing by adenosine deaminases that act on RNA. *Annu. Rev. Biochem.* **71,** 817–846 (2002).
25. Shaner, N. C. *et al.* Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nature Biotechnol.* **22,** 1567–1572 (2004).
26. Knight, S. W. & Bass, B. L. The role of RNA editing by ADARs in RNAi. *Mol. Cell* **10,** 809–817 (2002).
27. Wang, D. *et al.* Somatic misexpression of germline P granules and enhanced RNA interference in retinoblastoma pathway mutants. *Nature* **436,** 593–597 (2005).
28. Tabara, H., Yigit, E., Siomi, H. & Mello, C. C. The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-box helicase to direct RNAi in *C. elegans*. *Cell* **109,** 861–871 (2002).
29. Ketting, R. F., Haverkamp, T. H., van Luenen, H. G. & Plasterk, R. H. Mut-7 of *C. elegans*, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. *Cell* **99,** 133–141 (1999).
30. Yigit, E. *et al.* Analysis of the *C. elegans* Argonaute family reveals that distinct Argonautes act sequentially during RNAi. *Cell* **127,** 747–757 (2006).
31. Meister, G. *et al.* Identification of novel argonaute-associated proteins. *Curr. Biol.* **15,** 2149–2155 (2005).
32. Dalmay, T., Horsefield, R., Braunstein, T. H. & Baulcombe, D. C. SDE3 encodes an RNA helicase required for post-transcriptional gene silencing in *Arabidopsis*. *EMBO J.* **20,** 2069–2078 (2001).
33. Tomari, Y. *et al.* RISC assembly defects in the *Drosophila* RNAi mutant armitage. *Cell* **116,** 831–841 (2004).
34. Labrador, M. *et al.* Protein encoding by both DNA strands. *Nature* **409,** 1000 (2001).
35. Dorn, R., Reuter, G. & Loewendorf, A. Transgene analysis proves mRNA trans-splicing at the complex *mod(mdg4)* locus in *Drosophila*. *Proc. Natl Acad. Sci. USA* **98,** 9724–9729 (2001).
36. Horiuchi, T., Giniger, E. & Aigaki, T. Alternative *trans*-splicing of constant and variable exons of a *Drosophila* axon guidance gene, *lola*. *Genes Dev.* **17,** 2496–2501 (2003).
37. Felix, M. A. Genomes: a helpful cousin for our favourite worm. *Curr. Biol.* **14,** R75–R77 (2004).
38. Obbard, D. J., Jiggins, F. M., Halligan, D. L. & Little, T. J. Natural selection drives extremely rapid evolution in antiviral RNAi genes. *Curr. Biol.* **16,** 580–585 (2006).
39. Wilkins, C. *et al.* RNA interference is an antiviral defence mechanism in *Caenorhabditis elegans*. *Nature* **436,** 1044–1047 (2005).
40. Schott, D. H., Cureton, D. K., Whelan, S. P. & Hunter, C. P. An antiviral role for the RNA interference machinery in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **102,** 18420–18424 (2005).
41. Stiernagle, T. in *WormBook* (ed. the *C. elegans* research community) doi:10.1895/wormbook.1.101. 1 ⟨http://www.wormbook.org⟩ (2006).
42. Sempere, L. F., Dubrovsky, E. B., Dubrovskaya, V. A., Berger, E. M. & Ambros, V. The expression of the *let-7* small regulatory RNA is controlled by ecdysone during metamorphosis in *Drosophila melanogaster*. *Dev. Biol.* **244,** 170–179 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank: S. Kennedy for advice and initiating the enhanced RNAi screen; M. Finney for advice on advanced PCR; C. Zhang for *mg441* Eri characterization; Y. Kohara, D. Thierry-Mieg and J. Thierry-Mieg for EST sequences and clones; the Mitani laboratory for deletion strains; the Caenorhabditis Genetics Center for strains; Ruvkun laboratory members for reading the manuscript and many discussions; the laboratories of J. Kaplan and F. Ausubel for discussions; and the Leukemia and Lymphoma Society, and EMBO for funding to S.E.J.F.

Author Information The *eri-6* cDNA sequences are deposited in GenBank under accession numbers FJ009006–FJ009009. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to G.R. (ruvkun@molbio.mgh.harvard.edu).

METHODS

Isolation of enhanced RNAi mutants. *unc-47::gfp* worms were mutagenized with ethyl methanesulphonate, and both the F_2 and F_3 generations were cultured on lawns of *E. coli* expressing dsRNA that targeted *gfp*. Mutagenized populations were then screened for rare individuals showing a marked decrease in the number of γ-aminobutyric acid (GABA)-ergic neurons visibly expressing GFP under the dissecting microscope, screening through approximately 50,000 haploid genomes. We identified 44 mutants that show an enhanced exogenous RNAi phenotype. To ensure isolated mutants were Eri and to test for generally enhanced RNAi, we cultured them on *E. coli* targeting endogenous genes that were expressed in a variety of tissues, asking that they be hypersensitive compared with *unc-47::gfp* worms to at least two of the following three testers: *dpy-13*, *hmr-1* and *lin-1*, which result in dumpy, embryonic lethal and sterile, and multivulva phenotypes, respectively. Thermosensitive sterility was assayed by placing L1 worms at 25 °C and scoring for unfertilized eggs in the next generation. The 31 non-sterile mutants were in at least five complementation groups.

Mapping, genetic and expression analyses. *eri-6*(*mg379*) and *eri-7*(*mg411*) were mapped using a highly polymorphic *C. elegans* isolate, CB4856. Recombinants generated with either outcrossed strains carrying the *mg379* or the *mg411* alleles and CB4856 were assayed for their RNAi hypersensitivity phenotype or their ability to silence a *rol-6* transgene (*mgIs30 [rol-6*(*su1006*)*::lin-14* 3′-UTR, *col-10::lacZ, lim-6::gfp]*), respectively, and subsequently genotyped for single nucleotide polymorphisms that exist between the wild-type and CB4856 strains. *mg411* was placed into a 3.8-map-unit interval on chromosome I using the transgene silencing phenotype, whereas *mg379* was placed into a 4.6-map-unit interval using Eri phenotype *lin-1*(*RNAi*). Next, *mgIs30* was used to identify genes that inhibit transgene silencing using feeding RNAi targeting the open reading frames in the mapped *mg411* interval. Two open reading frames tested positive: C41D11.1 and C41D11.7. Sequencing of *mg411* and *mg379* identified the following mutations: *mg411* A→T at position 33105 of C41D11, and *mg379* G→A at position 34755 of cosmid C41D11. Five additional alleles were obtained by sequencing the *eri-6/7* locus in mutants from the ethyl methanesulphonate Eri screen. An eighth, presumably spontaneous, allele (see Supplementary Discussion), *eri-6/7*(*mg441*), was isolated by mating a serendipitously obtained Eri strain in the laboratory with *eri-6/7* mutants to ask for non-complementation, followed by sequencing of the *eri-6/7* gene.

Rescue experiments were done with the C41D11 cosmid at 5 μg ml^{-1}, and PCR fragments at 5 μg ml^{-1}. Promoter::*gfp* or m*rfp* fusions, and the *trans*-splicing reporter construct, were injected at 10 μg ml^{-1}. Full-length ERI-6/7::mCherry fusion was injected at 5 μg ml^{-1}. DiO was used to identify dye-filling amphid sensory neurons.

Enhanced RNAi assays. *unc-22*, 23-bp ds-siRNA injections were done as previously described[43]. The ds-siRNA was injected at 5 mg ml^{-1}. The percentage of F_1 twitchers was scored in 330 μM levamisole. Feeding RNAi assays were done at 20 °C. For *lin-1*(*RNAi*) and *unc-73*(*RNAi*), starved L1 s were placed on *E. coli* expressing the dsRNA. The next generation was scored for percentage displaying Muv or Pvl (*lin-1*), or coiler (*unc-73*) phenotypes. For *cel-1*(*RNAi*) and *hmr-1*(*RNAi*), worms that were exposed to RNAi starting at the L1 stage were scored as adults for the production of viable progeny. In complementation assays, heterozygotes of *mg369/mg379* and *mg411/tm1887* were tested for an Eri phenotype based on hypersensitivity to dsRNA targeting *hmr-1*.

cDNA analysis, 3′- and 5′-RACE. RNA was isolated from mutant worms or worms exposed to RNAi for several generations, using TRI Reagent (Molecular Research Center). RNA was treated with TurboDNase (Ambion) and cDNA was made using the RETROscript kit (Ambion). For 3′-RACE, the First Choice RLM-RACE kit (Ambion) was used; for 5′-RACE the SMART RACE kit (Clontech) was used. 3′- and 5′-RACE products were cloned using Qiagen PCR cloning kit. RT–PCR on worms expressing the *trans*-splicing reporter was done by hand-picking GFP-expressing worms for RNA isolation. RT–PCR was done using a primer in *gfp* combined with primers in *eri-6*. No PCR products were obtained with these primers using genomic DNA of GFP-expressing worms.

Western blotting. To enhance transgene expression, nematodes carrying the *trans*-splicing reporter were subjected to *rde-1* or *mut-16* RNAi by feeding. This resulted in brighter and broader expression of GFP. GFP-expressing worms from two independent transgenic lines were hand-picked into sample buffer. Non-transgenic N2 worms served as a negative control. Electrophoresis and western blotting were done according to standard methods. A monoclonal antibody to GFP (Clontech) was used at 1:10,000. A horseradish peroxidase-conjugated goat anti-mouse antibody was used as a secondary antibody.

Northern blotting. Northern blotting for endo-siRNAs was done as previously described[42] (http://chronic.dartmouth.edu/VRA/ambroslab.html) using 30 μg of RNA in 12% denaturing polyacrylamide gels. Prehybridization and hybridization were done in 7% SDS, 0.2 M Na-phosphate pH 7.0 at 35 °C. StarFire probe sequences were as described[5].

43. Caplen, N. J. *et al.* Rescue of polyglutamine-mediated cytotoxicity by double-stranded RNA-mediated RNA interference. *Hum. Mol. Genet.* **11**, 175–184 (2002).

ARTICLES

Crystal structure of opsin in its G-protein-interacting conformation

Patrick Scheerer[1]*, Jung Hee Park[1]*, Peter W. Hildebrand[1], Yong Ju Kim[1], Norbert Krauß[2], Hui-Woog Choe[1,3], Klaus Peter Hofmann[1,4] & Oliver P. Ernst[1]

Opsin, the ligand-free form of the G-protein-coupled receptor rhodopsin, at low pH adopts a conformationally distinct, active G-protein-binding state known as Ops*. A synthetic peptide derived from the main binding site of the heterotrimeric G protein—the carboxy terminus of the α-subunit (GαCT)—stabilizes Ops*. Here we present the 3.2 Å crystal structure of the bovine Ops*–GαCT peptide complex. GαCT binds to a site in opsin that is opened by an outward tilt of transmembrane helix (TM) 6, a pairing of TM5 and TM6, and a restructured TM7–helix 8 kink. Contacts along the inner surface of TM5 and TM6 induce an α-helical conformation in GαCT with a C-terminal reverse turn. Main-chain carbonyl groups in the reverse turn constitute the centre of a hydrogen-bonded network, which links the two receptor regions containing the conserved E(D)RY and NPxxY(x)$_{5,6}$F motifs. On the basis of the Ops*–GαCT structure and known conformational changes in Gα, we discuss signal transfer from the receptor to the G protein nucleotide-binding site.

G-protein-coupled receptors (GPCRs), also known as seven-transmembrane receptors, are the largest superfamily of plasma membrane receptors[1]. Binding of extracellular ligands to GPCRs modulates their capacity to catalyse GDP–GTP exchange in heterotrimeric G proteins (Gαβγ), thereby regulating the intracellular level of secondary messengers. Rhodopsin is the photoreceptor of the vertebrate retinal rod cell and the eponym of the rhodopsin family of GPCRs with ~670 members in the human genome[2]. Rhodopsin consists of the apoprotein opsin and the chromophore 11-*cis*-retinal, which is bound by a protonated Schiff base to Lys 296 in the seventh transmembrane helix to stabilize the inactive receptor state. Photon absorption leads to *cis–trans* isomerization of the retinal, which triggers deprotonation of the retinylidene Schiff base linkage and formation of the active, G-protein-binding metarhodopsin II state[3,4]. Metarhodopsin II decays by hydrolysis of the retinylidene Schiff base and release of the photolysed all-*trans*-retinal from its binding pocket to generate ligand-free opsin. Under physiological conditions, fresh 11-*cis*-retinal is metabolically supplied and taken up to regenerate rhodopsin[3,5]. Before regeneration is complete, visual sensitivity is reduced owing to a background level of GTP-bound G protein which is maintained by a persistently active form of opsin[5]. *In vitro*, opsin can readily adopt inactive (Ops) and active (Ops*) conformations, and low pH and a synthetic peptide derived from the C terminus of Gα stabilize Ops* (ref. 6).

Besides the crystal structures of bovine and squid rhodopsin[7–10] and different photointermediates[11,12], the structures of two isoforms of the β-adrenergic receptor[13,14] are known. All of these GPCR structures contain a ligand bound in their binding site and lack a prominent tilt of the cytoplasmic half of TM6 out of the helical bundle, which is considered to be mandatory for G protein activation[15–17]. We recently solved the crystal structure of ligand-free opsin and found that it was quite different from the known GPCR structures[18]. The most prominent features were the activating movement of TM6 and accompanying rearrangements in the regions of the conserved E(D)RY and NPxxY(x)$_{5,6}$F motifs[3,4,19]. Here we report the structure of Ops* when it actively binds to the key G protein interaction site. Major sites for receptor interaction include the C termini of the Gα and Gγ subunits of transducin ($G_t\alpha\beta\gamma$, consisting of $G\alpha_t$, $G\beta_1$ and $G\gamma_1$), the cognate G protein of rhodopsin[20–22]. We stabilized Ops* with an 11 amino acid synthetic peptide (GαCT, amino acid sequence ILENLKDCGLF, $G\alpha_t$(340–350)K341L) derived from the extreme C terminus of the transducin $G\alpha_t$ subunit[23]. The crystal structure of the Ops*–GαCT complex provides insight into the structural changes involved in signal transfer from the receptor to the G protein.

Structure of Ops*–GαCT complex

Optimized extraction of the receptor from native rod cell disc membranes allowed the growth of Ops*–GαCT complex crystals. We used two different approaches: co-crystallization of the GαCT peptide with either opsin or photoactivated rhodopsin, respectively. In the latter case, the rhodopsin and GαCT mixture was illuminated directly before crystallization to produce active metarhodopsin II. Loss of all-*trans*-retinal from the metarhodopsin II–GαCT complex occurred during crystallization and resulted in an Ops*–GαCT complex. This was determined by the lack of electron density for retinal in the 3.2 Å crystal structure which was obtained by co-crystallization of GαCT and photoactivated rhodopsin (Fig. 1 and Supplementary Fig. 1). When solubilized ligand-free opsin was used as the starting material, Ops*–GαCT crystals were obtained which diffracted to 3.8 Å and showed, within the limits of resolution, the same structure and binding of GαCT (data not shown). For crystallization, data collection, structure determination and refinement statistics see Methods and Supplementary Table 1.

The structure of Ops*–GαCT comprises amino acids 1–326 of opsin (lacking 22 C-terminal amino acids which are not resolved, presumably because of their known high mobility), and 11 amino acids of the GαCT peptide with well defined electron density (Fig. 1 and Supplementary Figs 1a and 2). The Ops* model includes the

[1]Institut für Medizinische Physik und Biophysik (CC2), Charité – Universitätsmedizin Berlin, Charitéplatz 1, D-10117 Berlin, Germany. [2]Queen Mary, University of London, School of Biological and Chemical Sciences, London E1 4NS, UK. [3]Department of Chemistry, College of Natural Science, Chonbuk National University, 561-756 Chonju, South Korea. [4]Zentrum für Biophysik und Bioinformatik, Humboldt-Universität zu Berlin, Invalidenstrasse 42, D-10115 Berlin, Germany.
*These authors contributed equally to this work.

seven transmembrane helices connected by extracellular (E1–E3) and cytoplasmic (C1–C3) loops, and the cytoplasmic helix 8 (H8), which runs along the membrane. The structure of Ops*–GαCT bears resemblance to the structure of peptide-free opsin (free Ops*)[18], which was used as a search model to solve the structure by molecular replacement. Compared to the Ops* structure[18], the Ops*–GαCT structure presented here shows subtle but distinct changes in Ops* induced by GαCT binding (Supplementary Fig. 4).

Gross structural features which are comparable in free Ops* and Ops*–GαCT but are different from the 11-*cis*-retinal-bound dark-state rhodopsin include a short helical turn in loop C1 and rearrangements of loops C2 and C3 (Fig. 2). Compared with dark-state rhodopsin, the cytoplasmic half of TM6 is tilted outwards of the helical bundle by 6–7 Å with Trp 265 as the pivot point, in agreement with predictions of the active receptor state derived from electron paramagnetic resonance spectroscopy[15,17]. Furthermore, TM5 is longer, very straight and more inclined than in rhodopsin, with a resulting shift of the cytoplasmic end by 2–3 Å towards TM6. As a consequence, the cytoplasmic ends of TM5 and TM6 are closer and nearly parallel to each other, forming a cytoplasmic helical pair. In the extracellular half, the Ops*–GαCT structure also shows two openings between TM5 and TM6 and between TM1 and TM7, respectively (data not shown), which may act as gates for retinal[18]. Owing to the TM5 and TM6 rearrangement, a crevice is opened into which the GαCT peptide binds in an α-helical conformation with an open reverse turn, an α_L-type C-capping motif (C-cap; Fig. 2a and Supplementary Figs 2–6). Notably, the conformation of GαCT observed in the Ops*–GαCT crystal is almost identical to the NMR solution structure of a GαCT analogue and the parent $G\alpha_t(340–350)$ peptide in the receptor-bound state determined by transfer nuclear Overhauser effect experiments[24,25] (Supplementary Fig. 3). As seen in the superposition of rhodopsin and GαCT (as part of Ops*–GαCT), dipping of GαCT into the cytoplasmic part of the helical bundle is not possible for the inactive dark state of rhodopsin with the given orientation of GαCT (Fig. 2b and Supplementary Fig. 5b).

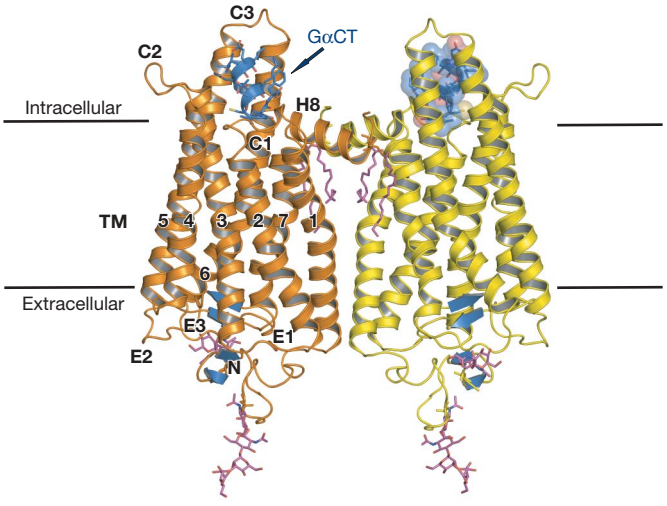


Figure 1 | Overall structure of Ops*–GαCT complex. The asymmetric unit contains one molecule of Ops* with one molecule of GαCT peptide bound at the intracellular side. Shown is a crystallographic dimer of symmetry related Ops*–GαCT monomers, with the view parallel to the membrane. The Ops* molecules (orange and yellow) and the GαCT peptides (blue) are shown in ribbon representation, and the GαCT side chains are shown as sticks. One of the GαCT molecules is also shown as transparent space-filling model. The GαCT peptide is derived from the $G\alpha_t$ C terminus ($G\alpha_t(340–350)$K341L). The K341L substitution increases the affinity for the active receptor conformation by two orders of magnitude[23]. Ops* consists of seven transmembrane helices (TM), connected by extracellular (E1–E3) and cytoplasmic (C1–C3) loops and the cytoplasmic helix H8. Two antiparallel β-sheets in the N terminus (strands β1 and β2) and loop E2 (strands β3 and β4) are shown as blue arrows. Oligosaccharides at Asn 2 and Asn 15 and palmitoyl chains at Cys 322 and Cys 323 are presented as sticks.

The dual role of Arg 135 in the E(D)RY motif

In rhodopsin, TM3 and TM6 are tethered by a hydrogen-bonded network ('ionic lock') which includes the side chains of Arg 135 and Glu 134 from the conserved E(D)RY motif in TM3, and the side chains of Glu 247 and Thr 251 in TM6 (Fig. 3a). In the peptide-free Ops* state[18] (Fig. 3b), as well as in Ops*–GαCT (Fig. 3c), the ionic lock is broken owing to the outward movement of TM6. The newly formed cytoplasmic TM5–TM6 pair conformation is stabilized by new interactions—that is, the side chains of Glu 247 and Thr 251 are released from Arg 135 and engage with the TM5 side chain of Lys 231 in a hydrogen-bonded network (not shown in Fig. 3b, c; see ref. 18). Apart from its main function in the ionic lock of rhodopsin, Arg 135—one of the most conserved residues in GPCRs—has a crucial role in the active receptor state. In both Ops* and Ops*–GαCT, the hydroxyl group of Tyr 223 from TM5 replaces the carboxyl group of Glu 134 by its interaction with Arg 135. The arginine side chain is thereby liberated from Glu 134 and allowed to swing into the centre of the GαCT binding crevice to form its floor. Stabilized by Tyr 223, it

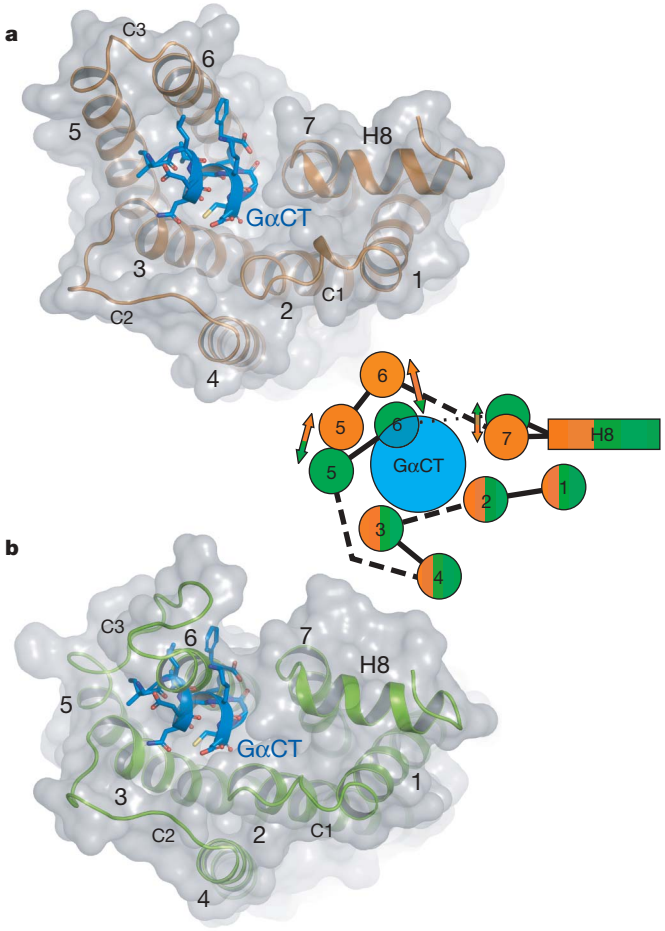


Figure 2 | Comparison of Ops*–GαCT and dark-state rhodopsin structures. **a**, View from the cytoplasm of Ops*–GαCT. Ops* is shown in ribbon (orange) and surface (grey) representation, the $G\alpha_t(340–350)$K341L peptide (ILENLKDCGLF) is shown as a ribbon and stick model (blue). The angle between the helical axis and the membrane normal (43 ± 2°) is in good agreement with predictions from residual dipolar coupling NMR data[25]. **b**, The view from the cytoplasm of dark-state rhodopsin (PDB accession 1U19; green ribbon and grey surface) and superposed GαCT (blue) is shown. The scheme indicates differences in the arrangement of TM1–TM7 between Ops*–GαCT (Ops*, orange; GαCT, blue) and dark-state rhodopsin (green).

then interacts with the backbone carbonyl group of Cys 347 at the tip of GαCT (Fig. 3c). Conversely, Glu 134 is not involved in an interaction with GαCT. Glu 134 faces towards TM2 and TM4, and further electron density is observed close to the carboxyl group of Glu 134. This electron density could be interpreted as a water molecule providing a weak linkage of Glu 134 to the main-chain carbonyl of Ala 153 and peptide nitrogen of Ile 75 (Supplementary Fig. 7). In this

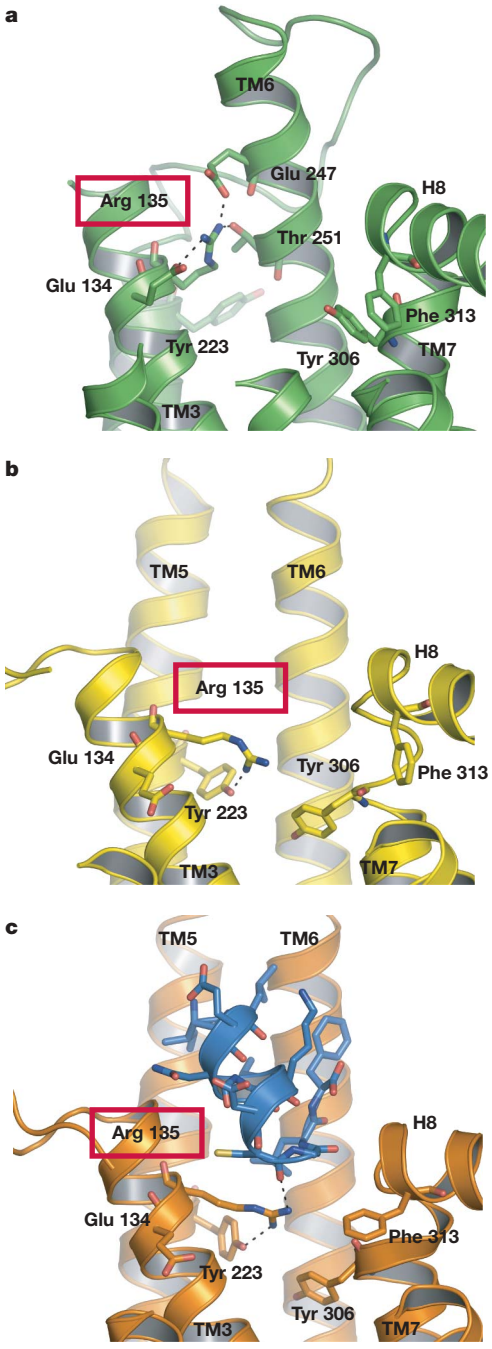


Figure 3 | Interactions of Arg 135 of the conserved E(D)RY motif. a, In dark-state rhodopsin (PDB accession 1U19), Arg 135 of the conserved E(D)RY motif interacts with Glu 134 and links TM3 with TM6 by Glu 247 and Thr 251. The TM7–H8 kink is stabilized by electrostatic interaction between Tyr 306 and Phe 313. **b**, In opsin (PDB accession 3CAP), Tyr 223 (TM5) interacts with Arg 135, which is released from Glu 134. Tyr 306 is rotated to face into the helical bundle. **c**, In Ops*–GαCT (Ops*, orange; GαCT, blue), Arg 135 and Tyr 223 interact with the main-chain carbonyl of GαCT residue Cys 347. In addition to the rotated Tyr 306 side chain, Phe 313 shows a different rotamer conformation. Parts of the helices in front were removed for clarity.

rather non-polar environment, Glu 134 is probably protonated which explains the proton uptake measured in the course of the light-activation of rhodopsin[4,26], which does not occur in the E134Q mutant[27]. The outwardly tilted position of TM6 and the new location of the Arg 135 side chain in Ops* and Ops*–GαCT is stabilized by Tyr 306 as part of the conserved NPxxY(x)$_{5,6}$F motif connecting TM7 and H8 (Fig. 3b, c). In rhodopsin, the side chains of Tyr 306 (TM7) and Phe 313 (H8) interact with each other, whereas in Ops* and Ops*–GαCT, Tyr 306 is extended into the helical bundle below Arg 135 and hinders the inward tilt of TM6.

GαCT and retinal-binding pockets

The inner side of the cytoplasmic TM5 and TM6 pair in Ops* provides a hydrophobic surface formed by Leu 226, Val 230 and Ala 233 in TM5, and by Thr 242, Thr 243, Ala 246 and Val 250 in TM6, for hydrophobic interaction with GαCT (Fig. 4a, Supplementary Figs 8 and 9 and Supplementary Table 2). Further hydrophobic contacts are provided by Val 138 and Val 139 at the junction of TM3 and loop C2, and by Leu 72 in loop C1. Some of the elements that contribute to this part of the binding site were predicted from earlier mapping approaches by mutational analysis[28,29] and photocrosslinking[30]. The fully induced C-capped helical structure of GαCT, which is not present in aqueous solution[24] or in GDP-bound $G_t\alpha\beta\gamma$[31,32], is used to connect the E(D)RY and NPxxY(x)$_{5,6}$F regions in Ops* (Fig. 4b). The main-chain carbonyl groups of Cys 347 and Lys 345 of GαCT are the bridge heads in an extended hydrogen-bonded network reaching from Cys 347 via Arg 135 (TM3) to Tyr 223 (TM5) on one side, and from Lys 345 via Gln 312 (H8) to Asn 310 (TM7) on the other side, explaining why mutations in the TM7–H8 kink affect binding of GαCT homologues[33]. For the Ops*-induced formation of the C-cap, the hydrogen bridges to Arg 135 and Gln 312 are probably the main determinants, together with GαCT-intramolecular hydrogen bonds. We assume that the C-cap is needed for precise recognition between GPCRs and G proteins. Contacts occur from fixed main chains and not flexible side chains (a mode of interaction known from transmembrane helices[34,35]), arguing for recognition of GαCT by means of the topology and specifics of local geometry.

Apart from local effects, GαCT binding also induces long-range stabilization effects into the ligand binding pocket. In Ops*, no electron density was observed for the side chain of Lys 296—the retinal attachment site. In Ops*–GαCT, however, electron density is observed, which indicates that the Lys 296 side chain is stabilized by a potential network of weak interactions between the ε-amino group of Lys 296 and the side chains of Ser 186 and Glu 181 in loop E2 (Fig. 4c). In this network, the carboxyl group of Glu 181 is also stabilized by the phenolic hydroxyl group of Tyr 268 (Fig. 4c, d). Notably, Glu 113—which is thought to form the counterion in inactive opsin[36]—is more than 6 Å from Lys 296 within Ops*–GαCT, which is too far to interact. Instead, the side chain of Glu 113 forms a hydrogen bridge with the main-chain nitrogen of Cys 187 in loop E2. Cys 187 in turn is further tethered to Cys 110 in TM3 by a conserved disulphide bridge. Concerning Lys 296, the resolution of the structure does not provide details of water in the changed hydrogen-bonded networks at the chromophore-binding site and the nearby NPxxY(x)$_{5,6}$F motif[8,9], so that not all the contributions to fixation of Lys 296 can be derived. However, it seems that G protein and ligand-binding sites are coupled in the active receptor conformation, fitting in with the classical receptor theory, in which the active conformation is stabilized by the ligand and/or G protein[37].

Ops* conformation in signal transduction and regeneration

The present analyses show that the same set of structural elements and conserved GPCR residues in the E(D)RY and NPxxY(x)$_{5,6}$F regions, which stabilize inactive rhodopsin, are used in different interactions in ligand-free Ops*, and are decisive in building new interactions of Ops* with GαCT. The same residues were crucial in previous analyses of the determinants of the metarhodopsin II

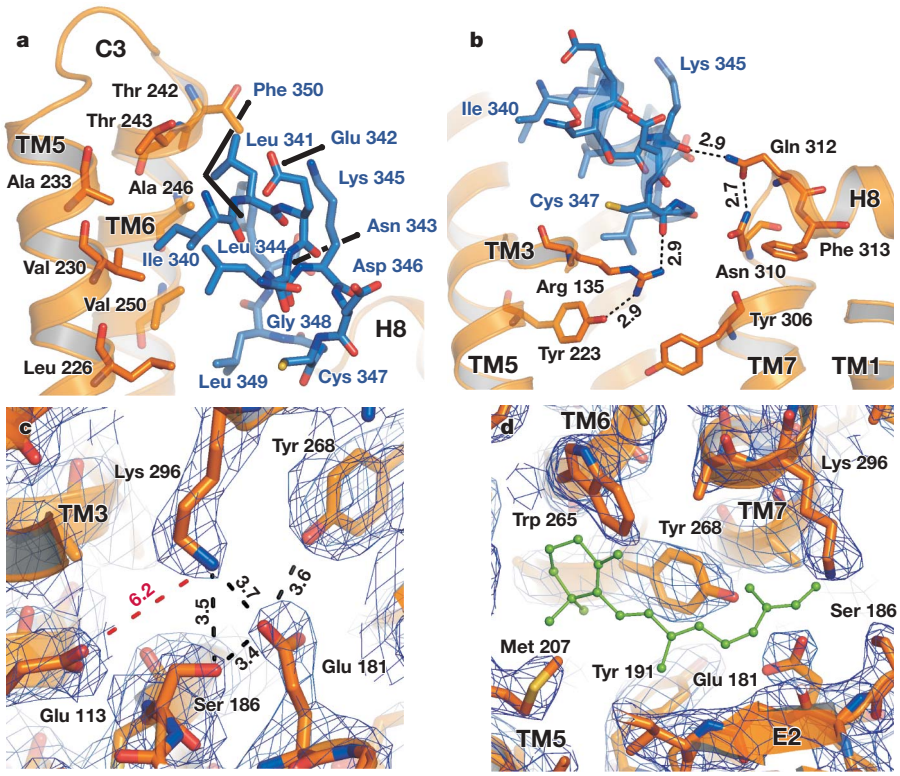


Figure 4 | Stabilizing effects of GαCT on Ops*. **a**, Pairing of the cytoplasmic parts of TM5 and TM6 in Ops* exposes hydrophobic side chains for van der Waals interaction with GαCT, which adopts a near-ideal α-helix terminated by an α_L-type C-cap. Amino acids of GαCT are labelled in blue font. **b**, The hydrogen-bonded network linking GαCT peptide with the two conserved E(D)RY (Arg 135) and NPxxY(x)$_{5,6}$F (Gln 312 and Asn 310) regions of Ops*. **c**, Long-range effects of GαCT binding lead to stabilization of the Lys 296 side chain in a potential network of weak interactions comprising Lys 296 in TM7, Ser 186 and Glu 181 in loop E2, and Tyr 268 in TM6. A direct stabilizing effect on Lys 296 by Glu 113 is negligible owing to a large distance of 6.2 Å between the side-chain amino and carboxyl groups. The numbers denote distances in Å (**b**, **c**). **d**, The view of the retinal binding pocket from the extracellular side is shown. 11-*cis*-retinal was superimposed from rhodopsin (PDB accession 1U19) to illustrate the empty retinal binding site. The side chains of Tyr 268 and Trp 265 moved into the space filled by retinal in rhodopsin. The blue mesh represents $2F_o - F_c$ electron density contoured at 1.0σ (**c**, **d**). Parts of the helices in front were removed for clarity.

state[19,27,38,39], with the hallmark of TM6 motion[4,15,17]. The structural information thus provides insights into how the Ops* state can act as the principal element in two different functional modules of the rod cell, that is, the signal transduction and the rhodopsin regeneration modules[40]. In signal transduction, the Ops* conformation is present in the highly active metarhodopsin II photoproduct, in which the covalently bound all-*trans*-retinal agonist shifts the pK for the transition to the active conformation into the neutral pH range[3,41]. In rhodopsin regeneration, Ops* facilitates the uptake of 11-*cis*-retinal into the binding site[42] and causes the visual system to behave as if it is experiencing a phenomenon equivalent to light ('equivalent background light'; ref. 5) owing to its capability—like metarhodopsin II—to activate the G protein.

A conceptual model for signal transfer to the G protein

We next studied whether the Ops*–GαCT structure provides a clue as to how the signal propagates from the receptor–GαCT interface to the 35–40 Å distant nucleotide-binding site in the G protein for GDP–GTP exchange. The Ops*–GαCT structure shows a well-defined interaction of a short fragment of $G\alpha_t$ with the receptor, so that the position of the $G_t\alpha\beta\gamma$ holoprotein in relation to the receptor and to the membrane can be determined. Therefore, we first modelled the $G\alpha_t$ C-terminal α5 helix as part of the Ops*–GαCT complex. For this, GαCT in Ops*–GαCT was amino-terminally elongated using the standard geometry of an α-helix (Supplementary Fig. 5). In a superposition of the resulting Ops*–GαCT and α5 model with the GDP-bound $G_t\alpha\beta\gamma$ crystal structure[31], $G_t\alpha\beta\gamma$ clashes with the lipid membrane (Fig. 5a). To avoid this, the $G_t\alpha\beta\gamma$ body (residues 1–324) has to be tilted relative to the helical C terminus (shown in red

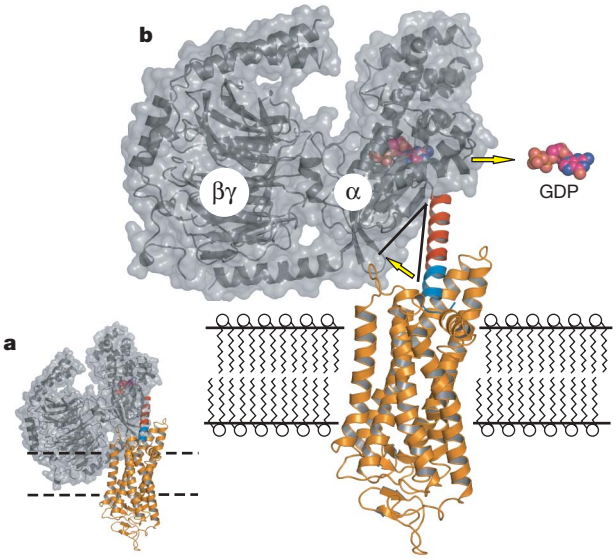


Figure 5 | Conceptual model for signal transmission from the active receptor to the G protein by the Gα C terminus. **a**, The Ops*–$G_t\alpha\beta\gamma$ complex modelled by superposition of $G_t\alpha\beta\gamma$ (PDB accession 1GOT; grey) and Ops*–GαCT (orange and blue, respectively) with a completed (red) C-terminal α5 helix (see Supplementary Fig. 5 for details of α5 helix modelling). **b**, To avoid a clash with the lipid bilayer, the remainder of $G_t\alpha\beta\gamma$ was tilted by about 40° relative to the α5 helix as indicated by the arrow. In this conformation, GDP is assumed to be released[43].

and blue in Fig. 5). A simple possibility is a change within the β6–α5 loop, which is involved in guanine-ring binding and connects the β6 strand ($G\alpha_t$ residues 316–320) with the C-terminal α5 helix and its C-cap ($G\alpha_t$ residues 325–350). A suitable rigid-body rotation and translation of the α5 helix relative to the $G_t\alpha\beta\gamma$ body on binding to activated rhodopsin was proposed from electron paramagnetic resonance studies on $G\alpha_{i1}$, a close homologue of $G\alpha_t$, and considered to be necessary for receptor-catalysed GDP release[22,43]. We may thus assume that the newly formed Ops* and GαCT interaction triggers GDP release using the α5 helix as a transmission rod, with extra constraints fixing the G protein relative to the receptor, to alter the β6–α5 loop conformation. An intact α5 helix was previously identified as a necessary element in this regard[43–45]. The model in Fig. 5b therefore probably represents a state in which the receptor has already induced emptying of the $G\alpha_t$ nucleotide-binding pocket[43].

Future work is needed to clarify the role of further structural elements involved in the receptor and G protein interaction. Although Gβγ is not mandatory for catalysed nucleotide exchange[46], modulation of the Gα–Gβγ interface in the holo G protein seems to facilitate GDP release (ref. 22 and references therein). A receptor-mimetic peptide derived from the N-terminal portion of loop C3 of the D_2-dopamine receptor activates the $G\alpha_i$ and $G\alpha_o$ subunits directly[47]. This peptide binds between the α4 helix and the β6 strand of $G\alpha_{i1}$–GDP, suggesting that the receptor loop C3 can contribute by means of the β6 strand to β6–α5 loop modulation[22,48]. It remains to be investigated how the active receptor recognizes the unstructured distal Gα C terminus and how an initial encounter/docking complex[49] between receptor and GDP-bound G protein is formed. A sequential interaction model has been proposed in which the G protein docks first by the farnesylated Gβγ subunit to the receptor to make the Gα C terminus available for receptor interaction[23,46].

METHODS SUMMARY

Native rhodopsin was purified by selective extraction from bovine rod outer-segment disc membranes using β-D-octylglucopyranoside. In a mixture with the GαCT peptide, rhodopsin was light-activated immediately before co-crystallization. Crystals grew within 10 days using hanging-drop vapour diffusion in a mixture of ammonium sulphate and MES or sodium citrate buffer at pH 6. Crystals were cryoprotected in 10% trehalose in precipitation buffer and frozen in liquid nitrogen for X-ray analysis at the synchrotron BESSY (Berlin, Germany). The Ops*–GαCT structure was solved by molecular replacement using the ligand-free opsin monomer (Protein Data Bank (PDB) accession 3CAP) as a search model.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 29 May; accepted 8 August 2008.

1. Pierce, K. L., Premont, R. T. & Lefkowitz, R. J. Seven-transmembrane receptors. *Nature Rev. Mol. Cell Biol.* **3,** 639–650 (2002).
2. Lagerstrom, M. C. & Schioth, H. B. Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nature Rev. Drug Discov.* **7,** 339–357 (2008).
3. Okada, T., Ernst, O. P., Palczewski, K. & Hofmann, K. P. Activation of rhodopsin: new insights from structural and biochemical studies. *Trends Biochem. Sci.* **26,** 318–324 (2001).
4. Knierim, B., Hofmann, K. P., Ernst, O. P. & Hubbell, W. L. Sequence of late molecular events in the activation of rhodopsin. *Proc. Natl Acad. Sci. USA* **104,** 20290–20295 (2007).
5. Lamb, T. D. & Pugh, E. N. Jr. Dark adaptation and the retinoid cycle of vision. *Prog. Retin. Eye Res.* **23,** 307–380 (2004).
6. Vogel, R. & Siebert, F. Conformations of the active and inactive states of opsin. *J. Biol. Chem.* **276,** 38487–38493 (2001).
7. Palczewski, K. *et al.* Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289,** 739–745 (2000).
8. Li, J., Edwards, P. C., Burghammer, M., Villa, C. & Schertler, G. F. Structure of bovine rhodopsin in a trigonal crystal form. *J. Mol. Biol.* **343,** 1409–1438 (2004).
9. Okada, T. *et al.* The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure. *J. Mol. Biol.* **342,** 571–583 (2004).
10. Murakami, M. & Kouyama, T. Crystal structure of squid rhodopsin. *Nature* **453,** 363–367 (2008).
11. Nakamichi, H. & Okada, T. Local peptide movement in the photoreaction intermediate of rhodopsin. *Proc. Natl Acad. Sci. USA* **103,** 12729–12734 (2006).
12. Salom, D. *et al.* Crystal structure of a photoactivated deprotonated intermediate of rhodopsin. *Proc. Natl Acad. Sci. USA* **103,** 16123–16128 (2006).
13. Warne, T. *et al.* Structure of a β_1-adrenergic G-protein-coupled receptor. *Nature* **454,** 486–491 (2008).
14. Cherezov, V. *et al.* High-resolution crystal structure of an engineered human β_2-adrenergic G protein-coupled receptor. *Science* **318,** 1258–1265 (2007).
15. Farrens, D. L., Altenbach, C., Yang, K., Hubbell, W. L. & Khorana, H. G. Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin. *Science* **274,** 768–770 (1996).
16. Sheikh, S. P., Zvyaga, T. A., Lichtarge, O., Sakmar, T. P. & Bourne, H. R. Rhodopsin activation blocked by metal-ion-binding sites linking transmembrane helices C and F. *Nature* **383,** 347–350 (1996).
17. Altenbach, C., Kusnetzow, A. K., Ernst, O. P., Hofmann, K. P. & Hubbell, W. L. High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation. *Proc. Natl Acad. Sci. USA* **105,** 7439–7444 (2008).
18. Park, J. H., Scheerer, P., Hofmann, K. P., Choe, H.-W. & Ernst, O. P. Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* **454,** 183–187 (2008).
19. Fritze, O. *et al.* Role of the conserved $NPxxY(x)_{5,6}F$ motif in the rhodopsin ground state and during activation. *Proc. Natl Acad. Sci. USA* **100,** 2290–2295 (2003).
20. Hamm, H. E. *et al.* Site of G protein binding to rhodopsin mapped with synthetic peptides from the α subunit. *Science* **241,** 832–835 (1988).
21. Kisselev, O. G., Ermolaeva, M. V. & Gautam, N. A farnesylated domain in the G protein γ subunit is a specific determinant of receptor coupling. *J. Biol. Chem.* **269,** 21399–21402 (1994).
22. Oldham, W. M. & Hamm, H. E. Heterotrimeric G protein activation by G-protein-coupled receptors. *Nature Rev. Mol. Cell Biol.* **9,** 60–71 (2008).
23. Herrmann, R. *et al.* Sequence of interactions in receptor-G protein coupling. *J. Biol. Chem.* **279,** 24283–24290 (2004).
24. Kisselev, O. G. *et al.* Light-activated rhodopsin induces structural binding motif in G protein α subunit. *Proc. Natl Acad. Sci. USA* **95,** 4270–4275 (1998).
25. Koenig, B. W. *et al.* Structure and orientation of a G protein fragment in the receptor bound state from residual dipolar couplings. *J. Mol. Biol.* **322,** 441–461 (2002).
26. Arnis, S. & Hofmann, K. P. Two different forms of metarhodopsin II: Schiff base deprotonation precedes proton uptake and signaling state. *Proc. Natl Acad. Sci. USA* **90,** 7849–7853 (1993).
27. Arnis, S., Fahmy, K., Hofmann, K. P. & Sakmar, T. P. A conserved carboxylic acid group mediates light-dependent proton uptake and signaling by rhodopsin. *J. Biol. Chem.* **269,** 23879–23881 (1994).
28. Acharya, S., Saad, Y. & Karnik, S. S. Transducin-α C-terminal peptide binding site consists of C-D and E-F loops of rhodopsin. *J. Biol. Chem.* **272,** 6519–6524 (1997).
29. Janz, J. M. & Farrens, D. L. Rhodopsin activation exposes a key hydrophobic binding site for the transducin α-subunit C terminus. *J. Biol. Chem.* **279,** 29767–29773 (2004).
30. Cai, K., Itoh, Y. & Khorana, H. G. Mapping of contact sites in complex formation between transducin and light-activated rhodopsin by covalent crosslinking: use of a photoactivatable reagent. *Proc. Natl Acad. Sci. USA* **98,** 4877–4882 (2001).
31. Lambright, D. G. *et al.* The 2.0 Å crystal structure of a heterotrimeric G protein. *Nature* **379,** 311–319 (1996).
32. Ridge, K. D. *et al.* Conformational changes associated with receptor stimulated guanine nucleotide exchange in a heterotrimeric G-protein α-subunit: NMR analysis of GTPγ S-bound states. *J. Biol. Chem.* **281,** 7635–7648 (2006).
33. Ernst, O. P. *et al.* Mutation of the fourth cytoplasmic loop of rhodopsin affects binding of transducin and peptides derived from the carboxyl-terminal sequences of transducin α and γ subunits. *J. Biol. Chem.* **275,** 1937–1943 (2000).
34. Edwards, M. D. *et al.* Pivotal role of the glycine-rich TM3 helix in gating the MscS mechanosensitive channel. *Nature Struct. Mol. Biol.* **12,** 113–119 (2005).
35. Hildebrand, P. W. *et al.* Hydrogen-bonding and packing features of membrane proteins: functional implications. *Biophys. J.* **94,** 1945–1953 (2008).
36. Cohen, G. B., Oprian, D. D. & Robinson, P. R. Mechanism of activation and inactivation of opsin: role of Glu 113 and Lys 296. *Biochemistry* **31,** 12592–12601 (1992).
37. De Lean, A., Stadel, J. M. & Lefkowitz, R. J. A ternary complex model explains the agonist-specific binding properties of the adenylate cyclase-coupled β-adrenergic receptor. *J. Biol. Chem.* **255,** 7108–7117 (1980).
38. Fahmy, K. & Sakmar, T. P. Regulation of the rhodopsin-transducin interaction by a highly conserved carboxylic acid group. *Biochemistry* **32,** 7229–7236 (1993).
39. Franke, R. R., König, B., Sakmar, T. P., Khorana, H. G. & Hofmann, K. P. Rhodopsin mutants that bind but fail to activate transducin. *Science* **250,** 123–125 (1990).
40. Hofmann, K. P., Spahn, C. M., Heinrich, R. & Heinemann, U. Building functional modules from molecular interactions. *Trends Biochem. Sci.* **31,** 497–508 (2006).
41. Meyer, C. K. *et al.* Signaling states of rhodopsin. Retinal provides a scaffold for activating proton transfer switches. *J. Biol. Chem.* **275,** 19713–19718 (2000).
42. Kefalov, V. J., Crouch, R. K. & Cornwall, M. C. Role of noncovalent binding of 11-*cis*-retinal to opsin in dark adaptation of rod and cone photoreceptors. *Neuron* **29,** 749–755 (2001).
43. Oldham, W. M., Van Eps, N., Preininger, A. M., Hubbell, W. L. & Hamm, H. E. Mechanism of the receptor-catalyzed activation of heterotrimeric G proteins. *Nature Struct. Mol. Biol.* **13,** 772–777 (2006).
44. Natochin, M., Moussaif, M. & Artemyev, N. O. Probing the mechanism of rhodopsin-catalyzed transducin activation. *J. Neurochem.* **77,** 202–210 (2001).

45. Marin, E. P., Krishna, A. G. & Sakmar, T. P. Disruption of the α5 helix of transducin impairs rhodopsin-catalyzed nucleotide exchange. *Biochemistry* **41,** 6988–6994 (2002).
46. Herrmann, R., Heck, M., Henklein, P., Hofmann, K. P. & Ernst, O. P. Signal transfer from GPCRs to G proteins: Role of the Gα N-terminal region in rhodopsin-transducin coupling. *J. Biol. Chem.* **281,** 30234–30241 (2006).
47. Nanoff, C. *et al.* The carboxyl terminus of the Gα-subunit is the latch for triggered activation of heterotrimeric G proteins. *Mol. Pharmacol.* **69,** 397–405 (2006).
48. Johnston, C. A. & Siderovski, D. P. Structural basis for nucleotide exchange on $G\alpha_i$ subunits and receptor coupling specificity. *Proc. Natl Acad. Sci. USA* **104,** 2001–2006 (2007).
49. Heck, M. & Hofmann, K. P. Maximal rate and nucleotide dependence of rhodopsin-catalyzed transducin activation: initial rate analysis based on a double displacement mechanism. *J. Biol. Chem.* **276,** 10000–10009 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Engelmann and C. Koch for technical assistance; P. Henklein for peptide synthesis; C. Enenkel and M. Sommer for critically reading the manuscript; U. Müller and the scientific staff of the Protein Structure Factory and the Freie Universität Berlin at beamlines BL 14.1 and BL 14.2 at BESSY for continuous support of the project. This work was supported by the Deutsche Forschungsgemeinschaft Sfb449 (to O.P.E.), Sfb740 (to O.P.E. and K.P.H.), DFG-KOSEF international cooperation ER 294/1-1 (to O.P.E.) and F01-2004-000-10054-0 (to H.-W.C.), and CBNU funds for overseas research 2006–2007 (to H.-W.C.) and a fellowship of the Leibniz Graduate School of Molecular Biophysics, Berlin (to Y.J.K.).

Author Information The atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession number 3DQB. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to O.P.E. (oliver.ernst@charite.de), K.P.H. (klaus_peter.hofmann@charite.de) or H.-W.C. (hwchoe@chonbuk.ac.kr).

METHODS

Peptide synthesis. Peptide GαCT (corresponding to $G\alpha_t$(340–350)K341L, amino acid sequence ILENLKDCGLF) was derived from the extreme C terminus of the $G\alpha_t$ subunit. It was synthesized with unmodified N and C termini as described[23,50]. The K341L substitution was found previously to increase the affinity of the peptide by two orders of magnitude compared with the native $G\alpha_t$(340–350) peptide[23,50].

Protein preparation and crystallization. Bovine rod outer segment membranes and opsin were prepared as described[18,51]. Rod outer segment membranes (10 mg ml^{-1} rhodopsin) and opsin (7 mg ml^{-1}) within rod outer segment membranes were solubilized essentially as described, using 1% β-D-octylglucopyranoside[9,52]. Synthetic GαCT peptide was added in a 4:1 molar ratio of peptide to solubilized rhodopsin. The mixture was incubated on ice for 1 h and illuminated for 15 s with green light (500 ± 20 nm). The illuminated sample was used without further purification for crystallization. Alternatively, GαCT was mixed with solubilized opsin in a 4:1 molar ratio for crystallization. Crystallization screens by the sparse matrix crystallization method[53] were carried out using the hanging-drop vapour diffusion method and more than 2,000 crystallization conditions. Promising conditions were systematically screened further by altering protein concentration, precipitation agents and pH. Ops*–GαCT crystals could be grown at 277 K using 24-well Linbro plates. Each hanging drop was prepared on a siliconized coverslip by mixing equal volumes (2 μl each) of receptor and GαCT mixture and reservoir solution. The reservoir solution contained 3.2 M ammonium sulphate in 0.1 M MES or sodium citrate buffer, pH 6.0. Crystals appeared within 5 days and continued to grow for 5 days. Fully grown crystals had dimensions of $0.05 \times 0.05 \times 0.15\ mm^3$.

Structure analysis. X-ray data collection of Ops*–GαCT crystals was performed at 100 K using cryoprotectant consisting of 90% (v/v) reservoir solution and 10% (w/v) trehalose. Diffraction data were collected at synchrotron beamline BL 14.2 (wavelength 0.9184Å) of the Protein Structure Factory and Freie Universität Berlin at BESSY with a MAR-165CCD detector. The crystal to detector distance was fixed at 240 mm. The rotation increment for each frame was 0.2° with an exposure time of 10 s. All images were indexed, integrated and scaled using HKL2000 (ref. 54). Ops*–GαCT crystals belong to rhombohedral space group *H*32 ($a = 238.02$ Å, $b = 238.02$ Å, $c = 109.13$ Å, $\alpha = \beta = 90°$, $\gamma = 120°$). Supplementary Table 1 summarizes the statistics for crystallographic data collection and structural refinement.

Initial phases for Ops*–GαCT were obtained by conventional molecular replacement protocol (rotation, translation and rigid body fitting) using the model of ligand-free opsin structure (PDB accession 3CAP) as an initial search trial. Molecular replacement was achieved using the CCP4 program PHASER[55] by first placing the seven transmembrane bundle of the opsin monomer (rotation function: $Z = 8.8$; translation function: $Z = 54.0$, as defined by PHASER). In subsequent steps, torsion angle molecular dynamics, simulated annealing using a slow-cooling protocol and a maximum likelihood target function, energy minimization, and B-factor refinement by the program CNS[56] were carried out in the resolution range 48.22–3.2 Å. After the first round of refinement, the GαCT peptide was clearly visible in the electron density of both σ_A-weighted $F_o - F_c$ maps, as well as in the simulated annealing omit density maps (as shown in Supplementary Fig. 2). Restrained individual B-factors were refined and the crystal structure was finalized by REFMAC5 and other programs in CCP4 (ref. 55). The final model had agreement factors R_{free} and R_{cryst} of 24.8% and 21.3%, respectively. Manual rebuilding of the opsin model and electron density interpretation were performed after each refinement cycle using the program COOT[57]. Structure validation was performed with the programs PROCHECK[58] and WHAT_CHECK[59]. Potential hydrogen bonds and van der Waals contacts (Supplementary Table 2) were analysed using the programs HBPLUS[60] and LIGPLOT[61] (Supplementary Fig. 8). All crystal structure superpositions of backbone alpha carbon traces were performed using CCP4 program LSQKAB (Figs 2b and 4d and Supplementary Figs 3b, c, 4a–c and 5)[55]. All molecular graphics representations were created using PyMol[62].

50. Herrmann, R. *et al.* Rhodopsin-transducin coupling: role of the Gα C-terminus in nucleotide exchange catalysis. *Vision Res.* **46,** 4582–4593 (2006).
51. Sachs, K., Maretzki, D. & Hofmann, K. P. Assays for activation of opsin by all-*trans*-retinal. *Methods Enzymol.* **315,** 238–251 (2000).
52. Murakami, M., Kitahara, R., Gotoh, T. & Kouyama, T. Crystallization and crystal properties of squid rhodopsin. *Acta Crystallogr. F* **63,** 475–479 (2007).
53. Jancarik, J. & Kim, S.-H. Sparse matrix sampling: a screening method for crystallization of proteins. *J. Appl. Crystallogr.* **24,** 409–411 (1991).
54. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326 (1997).
55. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50,** 760–763 (1994).
56. Brunger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).
57. Emsley, P. & Cowtan, K. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallogr. D* **60,** 2126–2132 (2004).
58. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26,** 283–291 (1993).
59. Hooft, R. W., Vriend, G., Sander, C. & Abola, E. E. Errors in protein structures. *Nature* **381,** 272 (1996).
60. McDonald, I. K. & Thornton, J. M. Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* **238,** 777–793 (1994).
61. Wallace, A. C., Laskowski, R. A. & Thornton, J. M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **8,** 127–134 (1995).
62. DeLano, W. L. The PyMOL Molecular Graphics System. <http://www.pymol.org> (2002).

LETTERS

Very fast optical flaring from a possible new Galactic magnetar

A. Stefanescu[1], G. Kanbach[1], A. Słowikowska[2,3], J. Greiner[1], S. McBreen[1] & G. Sala[1]

Highly luminous rapid flares are characteristic of processes around compact objects like white dwarfs, neutron stars and black holes. In the high-energy regime of X-rays and γ-rays, outbursts with variabilities on timescales of seconds or less are routinely observed, for example in γ-ray bursts[1] or soft γ-ray repeaters[2]. At optical wavelengths, flaring activity on such timescales has not been observed, other than from the prompt phase of one exceptional γ-ray burst[3]. This is mostly due to the fact that outbursts with strong, fast flaring are usually discovered in the high-energy regime; most optical follow-up observations of such transients use instruments with integration times exceeding tens of seconds, which are therefore unable to resolve fast variability. Here we show the observation of extremely bright and rapid optical flaring in the Galactic transient[4–7] SWIFT J195509.6+261406. Our optical light curves are phenomenologically similar to high-energy light curves of soft γ-ray repeaters and anomalous X-ray pulsars[8], which are thought to be neutron stars with extremely high magnetic fields (magnetars). This suggests that similar processes are in operation, but with strong emission in the optical, unlike in the case of other known magnetars.

SWIFT J195509.6+261406 was discovered as the γ-ray burst GRB 070610 (ref. 9) using the Burst Alert Telescope (BAT)[10] on board the NASA Swift spacecraft[11]. Subsequent observations in the X-ray (using Swift's X-ray telescope (XRT)[12] and the NASA Chandra X-ray Observatory[13]) and the optical bands have shown a point source compatible with the BAT error circle. The optical[4–6,14] and X-ray afterglow behaviour indicates that the source was probably not a γ-ray burst, but rather a Galactic X-ray transient[6,15]. It was therefore re-assigned the name SWIFT J195509.6+261406 (hereafter SWIFT J1955)[15].

Optical observations began just 421 s after the BAT trigger, using the OPTIMA-Burst[16] photo-polarimeter at the 1.3-m telescope of the Skinakas Observatory, Crete. A total of ~8.5 h of data was obtained during the five nights after the burst (Fig. 1). The overall optical light

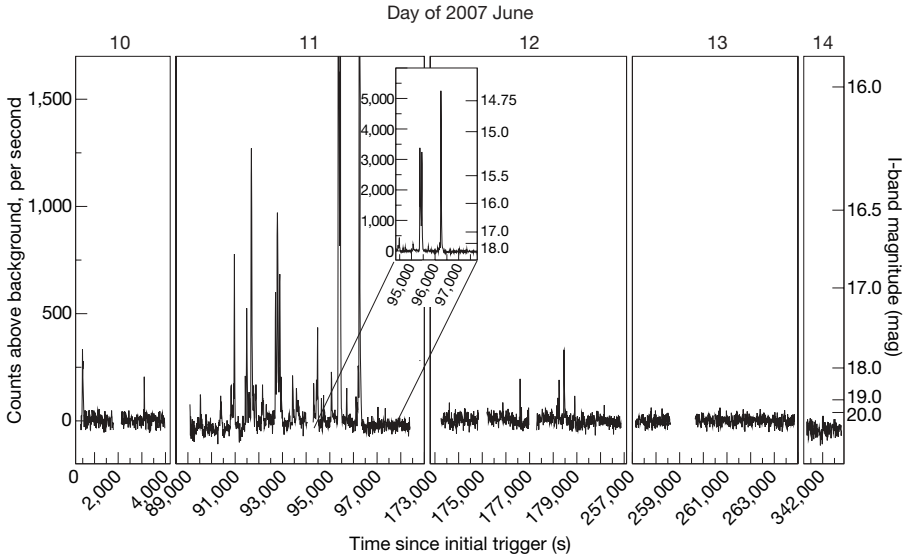


Figure 1 | Overview of overall optical high-time-resolution light curves of SWIFT J1955. The observations were obtained with OPTIMA-Burst[16], mounted at the 1.3-m Telescope of the Skinakas Observatory, Crete. OPTIMA-Burst is a fibre-fed system using six fibre apertures in a hexagonal bundle around the target fibre, plus one additional, more distant background fibre to determine and subtract the sky background. All apertures are of 300 μm diameter, corresponding to 6 arcsec on the sky with the SKO 1.3m telescope. The photon-counting mode of OPTIMA-Burst has an intrinsic photon-arrival-time resolution of 4 μs and records unfiltered white light radiation in the wavelength range 450–900 nm, with peak efficiency around 700 nm. To achieve a high signal-to-noise ratio, the recorded photon arrival times are binned into 10-s bins in the overview light curve shown here. The two most prominent flares, at 95,400 s and 96,250 s post trigger, are too bright to be shown adequately next to the rest of the flares, and are therefore displayed on a different scale in the inset (same axes as main panels). Owing to the very crowded field of SWIFT J1955, most of the available background channels were severely contaminated by field stars. The three least polluted background channels were used for background subtraction, but owing to changes in observing conditions from night to night, slight shifts (at the per cent level) in the zero-level occur between observation epochs. The observations were taken during a period of mediocre seeing, ranging from 1.5 to 2 arcsec. Simultaneous I-band (~630–1010-nm) observations, obtained with the IAC80 telescope of the Instituto de Astrofísica de Canarias (de Ugarte Postigo and Castro-Tirado, personal communication), were used to calibrate count rate to magnitudes.

[1]Max-Planck-Institute for Extraterrestrial Physics, PO Box 1312, 85741 Garching, Germany. [2]IESL, Foundation for Research and Technology - Hellas, PO Box 1385, GR-711 10 Heraklion, Greece. [3]Copernicus Astronomical Center, Rabiańska 8, 87-100 Toruń, Poland.

curve shows two brief flares shortly after the high-energy trigger on the first night[4] (2007 June 10). The second night shows a marked increase in flaring activity, culminating in two extremely bright flares with complex substructures[14] (2007 June 11). The morphology of the light curve, that is, isolated, brief bursts consisting of just one short peak and a bunching of these bursts in spurts of activity, is reminiscent of major X-ray outbursts in soft γ-ray repeaters (SGRs)[17].

The two most prominent flares, recorded near the end of the build-up in activity on 2007 June 11, are shown in detail in Fig. 2. A notable feature of these light curves is a very bright flare with an extremely steep rise followed by a slower, exponential decay. Superimposed on both flares are a number of secondary flares, again with a distinct FRED shape. The brightness on the rising edge changes by a factor of more than 200 over only 4 s, and possibly by even more than a factor of 10,000 when taking into account independent upper limits on the quiescent emission[7,18].

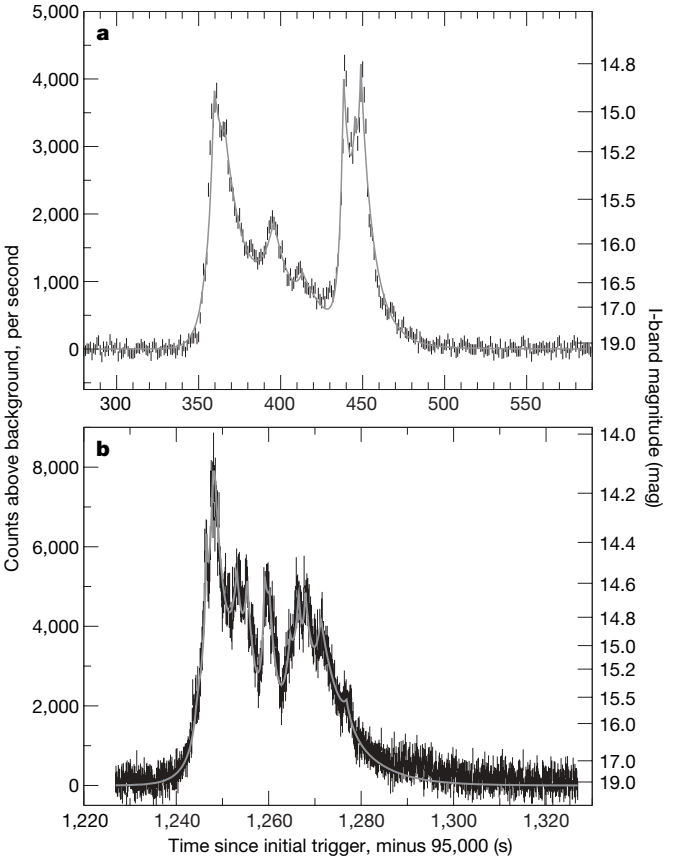


Figure 2 | Detailed light curves of the two most prominent flares of the epoch 2007 June 11 of Fig. 1. The data shown in **a** are binned to 1-s resolution and those in **b** to 0.1-s resolution to achieve reasonable signal-to-noise ratios but also resolve the fastest variabilities detected in the data. The error bars are 1σ statistical errors. Extremely fast variability is clearly visible and well resolved. The solid line in each panel is the result of a simple model, consisting of a superposition of several fast-rise, exponential-decay (FRED) subflares. The FRED shape is modelled by an exponential rise followed by an exponential decay, typically with a longer time constant in the exponential decay. The rise-time constants of the FRED curves fitted to the light curve in **a** range from 2.5 to 10 s, the decay-time constants range from 3 to 26 s. In **b**, the rise-time constants fitted to the light curve range from 0.4 to 3 s, and the decay timescales range from 0.3 to 7 s. The two flares are very bright, with peak I-band magnitudes of 14.8 mag (**a**) and 14.2 mag (**b**). The upper limit on the quiescent emission can be determined to be >20 mag (I band) using our data, and has been independently found to be >23.5 mag (I_c band) using Very Large Telescope data[18] or >24.5 mag (i′ band) using the Hale Telescope[7]. This means that the source changed brightness by a factor of at least 200 in the 4 s it took to go from the 1σ level to maximum brightness. If the quiescent emission was the same as in the Very Large Telescope and Hale Telescope data, the source possibly brightened by a factor $>10^4$.

The high intrinsic time-resolution of OPTIMA-Burst (4 μs) and the signal-to-noise ratio of the brightest parts of the light curve means that we can resolve the brightest features with a time resolution of ~10 ms. This enables us to fit a simple model of several superimposed FRED curves to the light curve segments shown in Fig. 2. The shortest timescales found in this analysis are 0.3–0.4 s. This places a limit on the maximum size of the emitting region of $\sim 10^{10}$ cm (about one-tenth of the Solar radius), because the light travel time across a larger region would blur the observed features.

To judge the luminosity and estimate the total optical-region energy output of the flares shown in Fig. 2, information about the source distance is required. Using X-ray extinction, plus optical, near infrared and radio observations, the distance to SWIFT J1955 can be estimated to be greater than 2–4 kpc, with a most probable distance of 4–8 kpc (ref. 18). In this work, we adopt a reference distance of 5 kpc.

Assuming a distance d (and that $A_V = 5$), the flares have maximum extinction-corrected I-band isotropic luminosities of $1.0 \times 10^{35}(d/5\,\mathrm{kpc})^2\,\mathrm{erg\,s^{-1}}$ (Fig. 2a) and $1.8 \times 10^{35}(d/5\,\mathrm{kpc})^2\,\mathrm{erg\,s^{-1}}$ (Fig. 2b), and total emitted energies in the I-band of $4.6 \times 10^{36}(d/5\,\mathrm{kpc})^2$ erg (Fig. 2a) and $3.2 \times 10^{36}(d/5\,\mathrm{kpc})^2$ erg (Fig. 2b). Taking into account the aforementioned size of the emitting region and putting the source at a distance of 5 kpc, a black-body temperature of several 10^7 K is necessary to explain the observed I-band flux in terms of thermal radiation. However, the extreme contrast ratio between quiescent emission and peak brightness, and the very short rise and decay times observed in these transitions, make it more probable that a non-thermal process is the source of the observed flares.

One possible non-thermal explanation is that the source is a magnetar, similar to an anomalous X-ray pulsar (AXP) or SGR. Optical and near-infrared emission has been observed previously in AXPs[19,20] and one SGR[21], but not with flaring variability. Considering the similarities between the optical light curve observed from SWIFT J1955 and X-ray light curves of SGR outbursts[17], it is conceivable that these are observations of optical flares in a magnetar. It is surprising that the strong X-ray flaring activity usually observed from such sources was not detected in this case. There were no simultaneous high-energy/optical observations during any observed period of flaring activity, so the amount of correlated X-ray/optical activity can only be assessed indirectly. The optical flaring activity during the first and second nights of observation (2 flares per hour and >7 flares per hour, respectively) contrasts significantly with the X-ray activity detected by XRT[7,18] between these two epochs. During the 5.5 h observed using the XRT, 10–40 flares would be expected if there were a strict correlation between optical and X-ray emission (which is not necessarily so, as in the case of the AXP XTE J1810-197 (ref. 22)) and the activity of the source were the same. However, only one significant X-ray flare, of 35-s duration, was observed by the XRT.

Because most of the observational data on magnetars so far gathered has been in the high-energy regime, the implications for optical emission have been somewhat neglected in theoretical works on magnetars. One intriguing model of an optical magnetar proposes optical ion cyclotron emission[23]. In this model, coherent microwave and radio emission emitted near the neutron star is absorbed higher in the magnetosphere by ions at their cyclotron resonance, and then re-emitted in the optical nearer to the poles, where the ion cyclotron cooling and transit times become comparable.

The power spectral densities of the two very bright flares (Fig. 2) are shown in Fig. 3. The power spectrum remains flat in the range between a few hertz and 1 kHz, which confirms very low variability on timescales shorter than ~0.1 s. Both power spectral densities show features similar to those in the X-ray regime that are called quasi-periodic oscillations. Although these features stem from only few oscillations and are therefore only moderately significant, it is notable that the features at a frequency of 0.16 ± 0.02 Hz (Fig. 3a) and

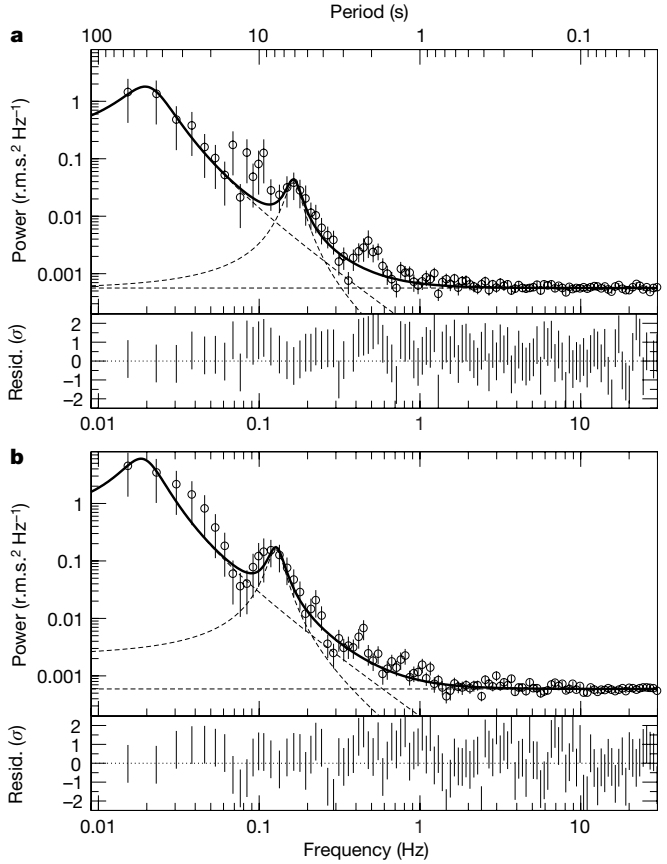


Figure 3 | Power spectral densities of the two prominent flares shown in Fig. 2. Each power spectrum is normalized such that its integral gives the squared root-mean-square (r.m.s.) fractional variability (expressed by giving each spectrum in units of r.m.s.^{2} Hz^{-1}). The error bars are 1σ statistical errors, calculated using the 'powspec' task of the Xronos timing-analysis software package (http://heasarc.gsfc.nasa.gov/docs/xanadu/xronos/). In each panel, the solid line is the result of a best-fit model consisting of two Lorentzians and a constant component. The dashed lines represent the individual components. In **a**, the reduced χ^2 statistic (sum of squared residuals divided by degrees of freedom) is 0.88, which is somewhat lower than desirable. The prominent Lorentzian peak in this panel has a central frequency of 0.163 ± 0.007 Hz and a full-width at half-maximum (FWHM) of 0.020 ± 0.009 Hz. In **b**, the reduced χ^2 statistic is 0.98. The central frequency of the peak in this panel is 0.127 ± 0.006 Hz and the FWHM is 0.015 ± 0.007 Hz. Note that the higher frequency Lorentzians in **a** and **b** have central frequencies that overlap within their respective FWHMs. Both power spectra stay flat in the frequency range between a few hertz and 1 kHz (not plotted here), confirming the limit on the fastest variability timescales from the multi-component fits. Resid., residual of the fit.

0.13 ± 0.02 Hz (Fig. 3b) overlap within their FWHMs. These features also coincide with a suggestion of possible X-ray periodicity reported at a frequency of 0.1446 Hz (ref. 7). The period of 6–8 s indicated by these observations lies in the range of typical rotational periods for AXPs[8] and SGRs[2]. If this is the rotational period of the magnetar, the radius of its light cylinder is about 3×10^{10} cm. This is consistent with the assumption of magnetospheric emission, as the upper limit for the size of the emitting region ($\sim 10^{10}$ cm) is less than the size of the light cylinder.

Two groups have proposed a similarity between SWIFT J1955 and the black hole X-ray binary V4641 Sgr[7,15]. However, the optical properties described here are quite different from those reported for V4641 Sgr. The light curve of SWIFT J1955 is dominated by variability that is extremely bright (>5 mag) and very fast (from a few seconds to <1 s), whereas the optical variability of V4641 Sgr is much less extreme, with flares of ~1 mag and ~50-s length[24,25]. Despite the fact that these values were derived from charge-coupled-device-based observations with a shortest time resolution of 30 s to minutes, it is clear that SWIFT J1955 is much more extreme in the magnitude and timescales of its variability. The timing properties of the fast, bright flares of SWIFT J1955, as well as independent multi-wavelength arguments[18], suggest a connection between this transient and optical magnetars.

Received 29 January; accepted 28 July 2008.

1. Mészáros, P. Gamma ray bursts. *Rep. Prog. Phys.* **69,** 2259–2321 (2006).
2. Kouveliotou, C. in *From X-Ray Binaries to Gamma-Ray Bursts* (eds van den Heuvel, E. P., Kaper, L., Rol, E. & Wijers, R. A. M. J.) 413–423 (Astronomical Society of the Pacific, 2003).
3. Racusin, J. L. *et al.* Broadband observations of the naked-eye γ-ray burst GRB 080319B. *Nature* **455,** 183–188 (2008).
4. Stefanescu, A. *et al.* GRB 070610: OPTIMA-Burst high-time-resolution optical observations. *GCN Circ.* **6492** (2007).
5. de Ugarte Postigo, A., Castro-Tirado, A. J. & Aceituno, F. GRB 070610: Optical observations from OSN. *GCN Circ.* **6501** (2007).
6. Kann, D. A. *et al.* GRB 070610: TLS RRM sees flaring behaviour - Galactic transient? *GCN Circ.* **6505** (2007).
7. Kasliwal, M. M. *et al.* GRB070610: A curious Galactic transient. *Astrophys. J.* **678,** 1127–1135 (2008).
8. Kaspi, V. Recent progress on anomalous X-ray pulsars. *Astrophys. Space Sci.* **308,** 1–11 (2007).
9. Pagani, C. *et al.* GRB 070610: Swift detection of a burst. *GCN Circ.* **6489** (2007).
10. Barthelmy, S. D. *et al.* The Burst Alert Telescope (BAT) on the SWIFT Midex Mission. *Space Sci. Rev.* **120,** 143–164 (2005).
11. Gehrels, N. *et al.* The Swift Gamma-Ray Burst Mission. *Astrophys. J.* **611,** 1005–1020 (2005).
12. Burrows, D. N. *et al.* The Swift X-Ray Telescope. *Space Sci. Rev.* **120,** 165–195 (2005).
13. Weisskopf, M. C. *et al.* An overview of the performance and scientific results from the Chandra X-Ray Observatory. *Publ. Astron. Soc. Pacif.* **114,** 1–24 (2002).
14. Stefanescu, A. *et al.* GRB 070610: OPTIMA-Burst detection of continued strong flaring activity. *GCN Circ.* **6508** (2007).
15. Markwardt, C. B. *et al.* SWIFT J195509.6+261406 / GRB 070610: A potential Galactic transient. *Astronom. Telegr.* **1102** (2007).
16. Kanbach, G. *et al.* in *High Time Resolution Astrophysics* (eds Phelan, D., Ryan, O. & Shearer, A.) 153–169 (Springer, 2008).
17. Hurley, K. *et al.* Reactivation and precise interplanetary network localization of the soft gamma repeater SGR 1900+14. *Astrophys. J.* **510,** L107–L109 (1999).
18. Castro-Tirado, A. J. *et al.* Flares from a candidate Galactic magnetar suggest a missing link to dim isolated neutron stars. *Nature* doi:10.1038/nature07328 (this issue).
19. Kaspi, V. *et al.* A major soft gamma repeater-like outburst and rotation glitch in the no-longer-so-anomalous X-Ray Pulsar 1E 2259+586. *Astrophys. J.* **588,** L93–L96 (2003).
20. Durant, M. & van Kerkwijk, M. H. Multiwavelength variability of the magnetar 4U 0142+61. *Astrophys. J.* **652,** 576–583 (2006).
21. Kosugi, G., Ogasawara, R. & Terada, H. A variable infrared counterpart to the soft gamma-ray repeater SGR 1806–20. *Astrophys. J.* **623,** L125–L128 (2005).
22. Testa, V. *et al.* Adaptive optics, near-infrared observations of magnetars. *Astron. Astrophys.* **482,** 607–615 (2008).
23. Beloborodov, A. M. & Thompson, C. Corona of magnetars. *Astrophys. J.* **657,** 967–993 (2007).
24. Uemura, M. *et al.* Outburst of a black hole X-ray binary V4641 Sgr in 2004 July. *Info. Bull. Var. Stars* **5626,** 1 (2005).
25. Uemura, M. *et al.* Optical observation of the 2003 outburst of a black hole X-Ray binary, V4641 Sagittarii. *Publ. Astron. Soc. Jpn* **56,** 823–829 (2004).

Acknowledgements We thank the Skinakas Observatory for their support and allocation of telescope time, and acknowledge the allocation of Chandra DDT time. We thank F. Schrey, T. Kougentakis and G. Paterakis for technical support, A. de Ugarte Postigo for access to private data taken simultaneous to some of our observations and A. Castro-Tirado for discussions. Skinakas Observatory is a collaborative project of the University of Crete, the Foundation for Research and Technology - Hellas, and the Max-Planck-Institute for Extraterrestrial Physics. A. Stefanescu acknowledges support from OPTICON. A. Słowikowska acknowledges support of the European Union through a Marie Curie Transfer of Knowledge Fellowship within the Sixth Framework Programme. S.McB. acknowledges the support of the European Union through a Marie Curie Intra-European Fellowship within the Sixth Framework Programme. G.S. is supported through DLR.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A. Stefanescu (astefan@mpe.mpg.de) or G.K. (gok@mpe.mpg.de).

LETTERS

Flares from a candidate Galactic magnetar suggest a missing link to dim isolated neutron stars

A. J. Castro-Tirado[1], A. de Ugarte Postigo[1,2], J. Gorosabel[1], M. Jelínek[1], T. A. Fatkhullin[3], V. V. Sokolov[3], P. Ferrero[4], D. A. Kann[4], S. Klose[4], D. Sluse[5], M. Bremer[6], J. M. Winters[6], D. Nuernberger[2], D. Pérez-Ramírez[7,8], M. A. Guerrero[1], J. French[9], G. Melady[9], L. Hanlon[9], B. McBreen[9], K. Leventis[10], S. B. Markoff[10], S. Leon[11], A. Kraus[12], F. J. Aceituno[1], R. Cunniffe[1,13], P. Kubánek[1,14], S. Vítek[1], S. Schulze[4], A. C. Wilson[15], R. Hudec[16], M. Durant[17], J. M. González-Pérez[17], T. Shahbaz[17], S. Guziy[18], S. B. Pandey[19], L. Pavlenko[20], E. Sonbas[3,21], S. A. Trushkin[3], N. N. Bursov[3], N. A. Nizhelskij[3], C. Sánchez-Fernández[22] & L. Sabau-Graziati[23]

Magnetars[1] are young neutron stars with very strong magnetic fields of the order of 10^{14}–10^{15} G. They are detected in our Galaxy either as soft γ-ray repeaters or anomalous X-ray pulsars. Soft γ-ray repeaters are a rare type of γ-ray transient sources that are occasionally detected as bursters in the high-energy sky[2–4]. No optical counterpart to the γ-ray flares or the quiescent source has yet been identified. Here we report multi-wavelength observations of a puzzling source, SWIFT J195509+261406. We detected more than 40 flaring episodes in the optical band over a time span of three days, and a faint infrared flare 11 days later, after which the source returned to quiescence. Our radio observations confirm a Galactic nature and establish a lower distance limit of ~3.7 kpc. We suggest that SWIFT J195509+261406 could be an isolated magnetar whose bursting activity has been detected at optical wavelengths, and for which the long-term X-ray emission is short-lived. In this case, a new manifestation of magnetar activity has been recorded and we can consider SWIFT J195509+261406 to be a link between the 'persistent' soft γ-ray repeaters/anomalous X-ray pulsars and dim isolated neutron stars.

Following the detection[5] of GRB 070610 as a single peaked γ-ray burst (GRB) lasting about 4.6 s and its bizarre X-ray counterpart[6–8] (dubbed SWIFT J195509+261406), we mounted a multi-wavelength observing campaign (see Supplementary Information sections 1–4 for details). Our data were collected starting ~1 min after the burst trigger time. In the first three nights of our observations, the source displayed strong optical flaring activity[7–9]. This, together with the location of the source in the Galactic plane, supported the view that the source is hosted by the Milky Way[10], and we give strong evidence for this here.

The flares from SWIFT J195509+261406 had durations ranging from tens of seconds to a few minutes and flux amplitudes up to about 100 times the 'outburst' baseline flux (or $\geq 10^4$ times the quiescent state). After 13 June, the activity decayed abruptly (Fig. 1) and no further flares were seen until 22 June, when a late-time, lower-brightness flare was detected in the near-infrared using the 8.2-m Very Large Telescope. A late-time observation by the XMM-Newton spacecraft ~173 days after the burst failed to detect the source, imposing an upper limit (3σ) to any underlying X-ray flux of $L_X \leq 3.1 \times 10^{-14}$ erg cm^{-2} s^{-1} (0.2–10 keV).

Our ^{12}CO ($J = 1-0$) spectrum towards the SWIFT J195509+261406 source reveals a molecular cloud at ~30 km s^{-1}, which contributes ~50% to the total column density N(H) derived by Swift/XRT (see Supplementary Information section 5.1). In fact, the overall Galactic column density along the line of sight towards $(l^{II}, b^{II}) = (63.5°, -1.0°)$ is $N(H) = N(H\,I) + 2N(H_2) = (14.1 \pm 2.0) \times 10^{21}$ cm^{-2} which should be compared with the X-ray absorption column derived from the Swift/XRT data: $10^{+4}_{-3} \times 10^{21}$ cm^{-2} from this work, or[7] $7.2^{+3}_{-2} \times 10^{21}$ cm^{-2} (all quoted errors here being 3σ). Therefore we conclude that SWIFT J195509+261406 is located in the Galaxy and beyond this particular molecular cloud at a kinematic distance of $D \approx 3.7$ kpc from the Sun. This value is consistent with ~4 kpc derived from the 'red clump' method (see Supplementary Information 5.2). Hereafter, we consider a reference distance of 5 kpc.

To discern the nature of the source, we explored several possibilities. The first is that the source resembles the 'bursting pulsar' GRO J1744–28 (refs 11,12). However, Swift/BAT has not recorded any other γ-ray burst from SWIFT J195509+261406 after the initial one. A second possibility is based on the proposed similarity to the black hole candidate V4641 Sgr[13,14]. This black hole, orbiting an intermediate-mass companion (a B9 subgiant), was suggested as the first member of the 'fast X-ray novae' group[15], and it has been proposed that SWIFT J195509+261406 is a member of this class[7]. However, several lines of evidence indicate otherwise. First, the lack of further detections of the baseline (non-flaring) flux during the outburst phase at γ-ray (Swift/BAT), millimetre (during the outburst, <0.6 mJy, 3σ; Supplementary Information 2) and centimetre

[1]Instituto de Astrofísica de Andalucía del Consejo Superior de Investigaciones Científicas (IAA-CSIC), PO Box 03004, E-18080 Granada, Spain. [2]European Southern Observatory, Casilla 19001, Santiago 19, Chile. [3]Special Astrophysical Observatory of Russian Academy of Science (SAO-RAS), Nizhnij Arkhyz, Karachai-Cherkessia, 369167 Russia. [4]Thüringer Landessternwarte Tautenburg, Sternwarte 5, D-07778 Tautenburg, Germany. [5]Laboratoire d'Astrophysique, École Polytechnique Fédérale de Lausanne (EPFL) Observatoire, 1290 Sauverny, Switzerland. [6]Institute de Radioastronomie Millimétrique (IRAM), 300 rue de la Piscine, 38406 Saint Martin d'Héres, France. [7]Facultad de Ciencias Experimentales, Universidad de Jaén, Campus Las Lagunillas, E-23071 Jaén, Spain. [8]Department of Physics and Astronomy, The University of Leicester, Leicester, LE1 7RH, UK. [9]School of Physics, University College Dublin, Dublin 4, Ireland. [10]Astronomical Institute 'Anton Pannekoek', University of Amsterdam, 1098 SJ Amsterdam, The Netherlands. [11]Institute de Radioastronomie Millimétrique (IRAM), Avda. Divina Pastora 7, Núcleo Central, E-18012 Granada, Spain. [12]Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany. [13]Cork Institute of Technology, Rossa Avenue, Bishopstown (Cork), Ireland. [14]Universidad de Valencia, Edif. Institutos de Investigación (GACE-ICMOL), Campus de Paterna, E-46980 Paterna (Valencia), Spain. [15]Department of Astronomy, University of Texas, Austin, Texas 78712, USA. [16]Astronomical Institute of the Czech Academy of Sciences, Fričova 298, 25165 Ondřejov, Czech Republic. [17]Instituto de Astrofísica de Canarias (IAC), Via Láctea s/n, E-38205 La Laguna (Tenerife), Spain. [18]Nikolaev State University, Nikolskaya 24, 54030 Nikolaev, Ukraine. [19]Aryabhatta Research Institute of Observational-Sciences (ARIES), Manora Peak, NainiTal, Uttarakhand, 263129, India. [20]Crimean Astrophysical Observatory, 98409 Nauchny, Ukraine. [21]University of Cukurova, Department of Physics, 01330 Adana, Turkey. [22]European Space Astronomy Centre (ESAC), Avenida de los Castillos s/n, Urbanizacion Villafranca del Castillo, E-28691 Villanueva de la Cañada (Madrid), Spain. [23]Instituto Nacional de Técnica Aeroespacial (INTA), E-28750 Torrejón de Ardoz (Madrid), Spain.

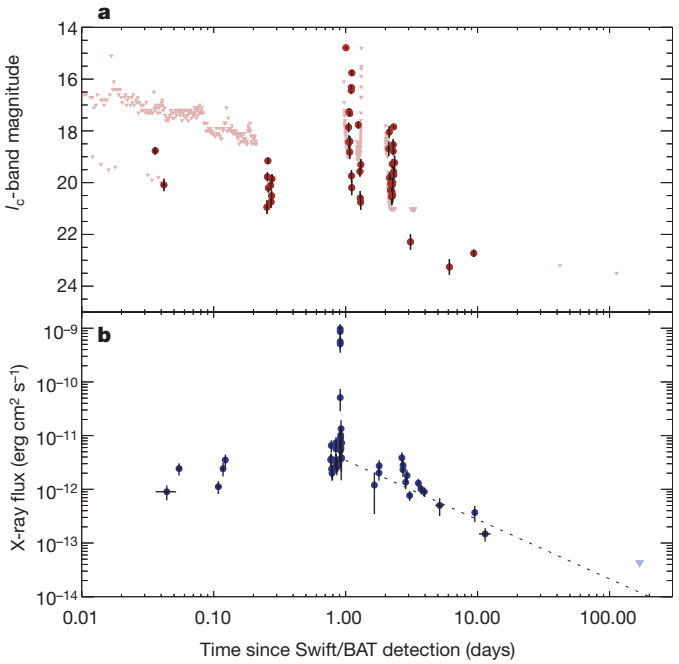


Figure 1 | Optical and X-ray light curves of SWIFT J195509+261406 (June–November 2007). **a**, Optical detections (I_c-band magnitudes, filled circles, with 1σ error bars) are shown together with 3σ upper limits (triangles). **b**, Swift X-ray data (0.2–10 keV, filled circles, with 1σ error bars) together with the late-time 3σ limit obtained with XMM-Newton (triangle). Both light curves show strong activity during the first three days, reaching the maximum around one day after the γ-ray burst and gradually decaying after the third day until the source became undetectable. The X-ray observations made by Swift do not overlap with the times of any of the optical flares that we have recorded. However, observations in both X-ray and optical agree that the strongest flaring activity is found around one day after the γ-ray event. A short (~30-s) powerful X-ray flare, for which the flux increased by a factor of $\Delta f/f \approx 100$ on a timescale of $\Delta t/t \approx 10^{-4}$, was followed by several optical flares of similar amplitude. The X-ray data one day after the giant flare event (excluding minor flaring-like activity) can be fitted by a power-law decay $F \propto t^{\alpha}$ with $\alpha = -0.75 \pm 0.25$, consistent with the values seen in the decline phase of the anomalous X-ray pulsar[22] XTE J1810-197 and the transient magnetar[23] SGR 1627-41. The baseline X-ray luminosity of SWIFT J195509+261406 during the first 8,000 s that followed the initial γ-ray spike was $\sim 1.2 \times 10^{34}\,(D/5\,\mathrm{kpc})^2\,\mathrm{erg\,s^{-1}}$ whereas the quiescent X-ray luminosity was $\leq 9.0 \times 10^{31}\,(D/5\,\mathrm{kpc})^2\,\mathrm{erg\,s^{-1}}$ (0.2–10 keV) at our late-time X-ray observation after ~173 days. This is in any case significantly smaller than the values of $\sim 4 \times 10^{33}\,\mathrm{erg\,s^{-1}}$ and $\sim 1.3 \times 10^{35}\,\mathrm{erg\,s^{-1}}$ derived for SGR 1627-41 and some anomalous X-ray pulsars, respectively.

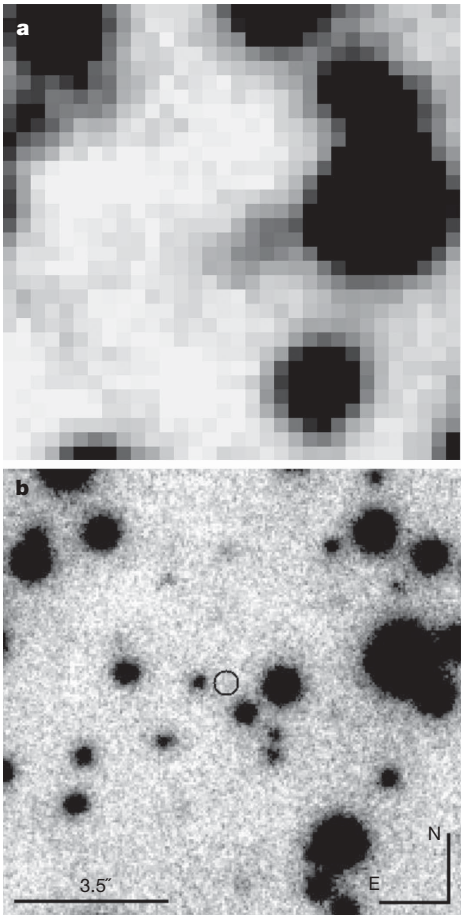


Figure 2 | Deep, late observations of the SWIFT J195509+261406 field. **a**, Deep I_c-band image obtained with the 6.0-m Big Telescope Altazimuthal (using SCORPIO) on 12 October 2007. **b**, Deep H-band image obtained on 30 September 2007 with the 8.2-m Very Large Telescope (using NACO) using natural guide-star adaptive optics. Both images show that the source has disappeared. The location for SWIFT J195509+261406 is marked with a circle (error radius of 0.26″). The limiting magnitudes are $I_c > 23.5$ and $H > 23.0$. There is no evidence for an underlying supernova remnant or a massive star cluster. Four anomalous X-ray pulsars, a subclass of magnetars, have been detected at near-infrared wavelengths[26] but no H-band counterpart of any SGR is known. SGR 1806–20, which is hidden by more than 30 mag of visual extinction, was only seen in the K-band when it was in an active state[27]. The other three known SGRs have no near-infrared counterparts.

wavelengths[7] (during the decline phase, < 0.09 mJy, 3σ) implies a different physical mechanism, because considerable γ-ray and radio emission (the latter arising from a collimated jet) was recorded at the time of the V4641 Sgr outbursts[16]. Moreover, the near-infrared limit imposed on the quiescent counterpart of SWIFT J195509+261406 ($H > 23$; that is, $M_H > 8.1$ assuming $D \approx 5$ kpc and $E(B-V) = 1.9 \pm 0.6$ towards the source; see Fig. 2 and Supplementary Information 5) constrains the spectral type of any companion to a main-sequence star with spectral type later than[17] M5V (that is, $\leq 0.12M$ solar masses) unless the donor was a hydrogen-poor (semi-) degenerate star in an ultra-compact X-ray binary[18] (with orbital period less than 1 h).

In fact, an ultra-compact low-mass X-ray binary with blobs of homogeneous synchrotron-emitting plasma of size $\sim 10^7$ cm and magnetic fields of strength $\sim 10^5$ G can explain both the observed baseline flux and the flaring episodes in X-ray and optical wavelengths. This third scheme allows such blobs to be found in a magnetized corona[19] (an extended region of low-density, X-ray irradiated material above and below the accretion disk) or a wind, rather than in the outer regions of a collimated jet. Thus, SWIFT J195509+261406 could be part of such a system.

On the other hand, we point to a final possibility: that the SWIFT source is an isolated compact object, that is, a new magnetar in our Galaxy, which displays activity like that of a soft γ-ray repeater (SGR) in the optical; and from which only one hard burst was recorded in γ-rays, near the onset of its bursting activity. If this is the case, SWIFT J195509+261406 becomes the first SGR detected at optical wavelengths. This would be supported by the burst durations (Fig. 3) and the timing properties of the flares[8].

How could the optical flares be produced in SWIFT J195509+261406? One possibility is, according to magnetar corona models[20], that the flares are due to coherent plasma bunches, and their luminosity depends on the unknown bunching factor, leading easily to $L_{\mathrm{opt}} \approx 10^{35}\,(D/5\,\mathrm{kpc})^2\,\mathrm{erg\,s^{-1}}$, as observed at optical frequencies[8].

In contrast to SGR 0526–66, SGR 1806–20 and SGR 1900+14, which all show 'persistent' X-ray emission in the range $\sim 10^{-11}$ to $10^{-12}\,\mathrm{erg\,cm^{-2}\,s^{-1}}$, SWIFT J195509+261406 strongly resembles the

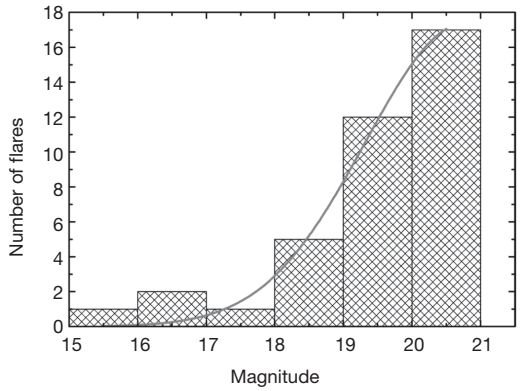


Figure 3 | Log-normal distribution of flare fluxes for SWIFT J195509+261406. The magnitude distribution of the optical flares detected from SWIFT J195509+261406 in the I_c band is shown. Using all I_c-band detections of the source, we find that the flare fluxes are log-normally distributed as seen in the high-energy flares of SGR 1806-20 (ref. 28) and SGR 1900+14 (ref. 29), supporting the claim that SWIFT J195509+261406 is a new SGR, although this is not conclusive. We find that the observed data are well fitted by a truncated normal distribution (solid line) with $N = \frac{A}{w\sqrt{\pi/2}} e^{-2(x-x_c)^2/w^2}$. Here, N is the number of flares in a one-magnitude bin, x is the magnitude, x_c is the centre of the distribution, w is the width and A is the amplitude. The fit is moderately acceptable with $\chi^2 = 49.7$ for 35 degrees of freedom, and we find as parameters $x_c = 20.80 \pm 0.61$ mag, $w = 2.96 \pm 0.76$ mag and $A = 64.5 \pm 22.7$. The truncation of the distribution is a natural result of the limiting magnitude of the observations. In addition, we note that intermediate-duration bursts (~1–30 s) have been recorded in SGR 1627-41 and SGR 1900+14. In particular, two events arising from SGR 1900+14 (at ~7 kpc) and lasting about 1 s displayed unusual hard power-law spectra[30] similar and comparable in energy, $E = (6.5–11) \times 10^{39}$ erg, to[5] GRB 070610, the burst of γ-rays associated with SWIFT J195509+261406 ($E = 1.9 \times 10^{39}\ (D/5\ \mathrm{kpc})^2$ erg). Thus, both the γ-ray burst duration and the log-normal distribution of the optical flares strengthen the association of the SWIFT source with a magnetar, although the lack of persistent X-ray pulses (which would allow us to determine the spin period derivative and thus the magnetic field) currently prevents us from proving the existence of extreme magnetic fields typical of magnetars, including SGRs and anomalous X-ray pulsars.

'transient' behaviour of SGR 1627–41 (ref. 21) and another magnetar, XTE J1810-197 (ref. 22). The former source experienced an activity period for 6 weeks in 1998 and its underlying X-ray flux was then observed to decrease by a factor of ~50 over a timespan of 1,000 days, flattening off at $\sim 3 \times 10^{-13}$ erg cm^{-2} s^{-1}, implying a quiescent X-ray luminosity[23] $\sim 4 \times 10^{33}$ erg s^{-1}, which is still significantly higher than the value of $\leq 9.0 \times 10^{31}\ (D/5\ \mathrm{kpc})^2$ erg s^{-1} derived for the SWIFT source from our late-time X-ray observation after ~173 days. Therefore we suggest that SWIFT J195509+261406 could be the first magnetar (either persistent or transient) in our Galaxy that shows strong and protracted optical flaring activity.

A deeper X-ray observation together with a detailed study of future activity periods of SWIFT J195509+261406, including simultaneous X-ray/optical monitoring, could shed light onto its nature and discern whether the source is an ultra-compact low-mass X-ray binary or an isolated neutron star displaying a new manifestation of magnetar activity. In the latter case, it would represent a link between 'persistent' SGRs/anomalous X-ray pulsars (with $L_X \approx (2–4) \times 10^{35}$ erg s^{-1} and $\sim(0.2–5) \times 10^{35}$ erg s^{-1} respectively) and dim isolated neutron stars[24] (with $L_X \approx (2–20) \times 10^{30}$ erg s^{-1}), being one of a few hundred Galactic magnetars to become active[25] in the past $\sim 10^4$ years.

Received 31 January; accepted 31 July 2008.

1. Thompson, C. & Duncan, R. C. The soft gamma repeaters as very strongly magnetized neutron stars. II. Quiescent neutrino, X-ray, and Alfvén wave emission. *Astrophys. J.* **473,** 322–342 (1996).
2. Mazets, E. P., Golenetskii, S. V., Gurian, Iu. A. & Ilinskii, V. N. The 5 March 1979 event and the distinct class of short gamma bursts: Are they of the same origin? *Astrophys. Space Sci.* **84,** 173–189 (1982).
3. Kouveliotou, C. *et al.* The rarity of soft gamma-ray repeaters deduced from reactivation of SGR 1806–20. *Nature* **368,** 125–127 (1994).
4. Hurley, K. in *Gamma-Ray Bursts, 5th Huntsville Symposium, Huntsville* (eds Kippen, R. M., Mallozzi, R. S. & Fishman, G. J.) *AIP Conf. Proc.* **526,** 763–770 (2000).
5. Tueller, J. *et al.* GRB 070610, Swift-BAT refined analysis. *GCN Circ.* 6491 (2007).
6. Pagani, C. & Kennea, J. A. GRB 070610; Swift-XRT position. *GCN Circ.* 6490 (2007).
7. Kasliwal, M. M. *et al.* GRB 070610: A curious galactic transient. *Astrophys. J.* **678,** 1127–1135 (2008).
8. Stefanescu, A. *et al.* Very fast optical flaring from a possible new Galactic magnetar. *Nature* doi:10.1038/nature07308 (this issue).
9. de Ugarte Postigo, A., Castro-Tirado, A. J. & Aceituno, F. GRB 070610: Optical observations from OSN. *GCN Circ.* 6501 (2007).
10. Kann, D. A. *et al.* GRB 070610: TLS RRM sees flaring behaviour: Galactic transient? *GCN Circ.* 6505 (2007).
11. Kouveliotou, C. *et al.* A new type of transient high-energy source in the direction of the Galactic Centre. *Nature* **379,** 799–801 (1996).
12. Sazonov, S., Sunyaev, R. & Lund, N. Super-Eddington X-ray luminosity of the bursting pulsar GRO J1744–28: WATCH/GRANAT observations. *Astron. Lett.* **23,** 326–334 (2005).
13. Markwardt, C. *et al.* SWIFT J195509.6+261406 / GRB 070610: A potential Galactic transient. *Astron. Tel.* 1102 (2007).
14. Revnivtsev, M., Gilfanov, M., Churazov, E. & Sunyaev, R. Super-Eddington outburst of V4641 Sgr. *Astron. Astrophys.* **391,** 1013–1022 (2002).
15. Uemura, M. *et al.* The 1999 optical outburst of the fast X-ray nova, V4641 Sagittarii. *Publ. Astron. Soc. Jpn* **54,** 95–101 (2002).
16. Hjellming, R. M. *et al.* Light curves and radio structure of the 1999 September transient event in V4641 Sagittarii (=XTE J1819–254=SAX J1819.3–2525). *Astrophys. J.* **544,** 977–992 (2000).
17. Dahn, C. C. *et al.* Astrometry and photometry for cool dwarfs and brown dwarfs. *Astron. J.* **124,** 1170–1189 (2002).
18. in't Zand, J. J. M., Jonker, P. G. & Markwardt, C. B. Six new candidate ultracompact X-ray binaries. *Astron. Astrophys.* **465,** 953–963 (2007).
19. Merloni, A., Di Matteo, T. & Fabian, A. C. Magnetic flares and the optical variability of the X-ray transient XTE J1118+480. *Mon. Not. R. Astron. Soc.* **318,** L15–L19 (2000).
20. Thompson, C. & Beloborodov, A. M. High-energy emission from magnetars. *Astrophys. J.* **634,** 565–569 (2005).
21. Woods, P. M. *et al.* Discovery of a new soft gamma-ray repeater, SGR1627–41. *Astrophys. J.* **519,** L139–L142 (1999).
22. Ibrahim, A. I. *et al.* Discovery of a transient magnetar: XTE J1810–197. *Astrophys. J.* **609,** L21–L24 (2004).
23. Mereghetti, S., Esposito, P. & Tiengo, A. XMM-Newton observations of soft gamma-ray repeaters. *Astrophys. Space Sci.* **308,** 13–23 (2007).
24. Treves, A., Turolla, R., Zana, S. & Colpi, M. Isolated neutron stars: Accretors and coolers. *Publ. Astron. Soc. Pacif.* **112,** 297–314 (2000).
25. Muno, M. P. *et al.* A search for new Galactic magnetars in archival Chandra and XMM-Newton observations. *Astrophys. J.* **680,** 639–653 (2008).
26. McGill Pulsar Group. SGR/AXP online catalogue. Available at http://www.physics.mcgill.ca/~pulsar/magnetar/main.html.
27. Israel, G. *et al.* Discovery and monitoring of the likely IR counterpart of SGR 1806–20 during the 2004 γ-ray burst-active state. *Astron. Astrophys.* **438,** L1–L4 (2005).
28. Hurley, K., McBreen, B., Delaney, M. & Britton, A. Lognormal properties of SGR 1806–20 and implications for other SGR sources. *Astrophys. Space Sci.* **231,** 81–84 (1995).
29. Göğüş, E. *et al.* Statistical properties of SGR 1900+14 bursts. *Astrophys. J.* **526,** L93–L96 (1999).
30. Woods, P. M. *et al.* Hard burst emission from the soft gamma repeater SGR 1900+14. *Astrophys. J.* **527,** L47–L50 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work is based on observations carried out with the 0.3-m robotic telescope at the Spanish BOOTES-2 astronomical station of the Estación Experimental de La Mayora (CSIC), the 0.4-m WATCHER telescope operated by UCD at Boyden Observatory (South Africa), the 0.8-m IAC telescope at the Spanish Observatorio de Izaña of the Instituto de Astrofísica de Canarias (IAC), the 1.2-m Mercator telescope operated by the Flemish Community at the Spanish Observatorio del Roque de los Muchachos of the IAC, the 1.34-m telescope at the Tautenburg Observatory (Germany), the 1.5-m OSN telescope at the Spanish Observatorio de Sierra Nevada of the Instituto de Astrofísica de Andalucía (CSIC), the 6.0-m BTA telescope at the Special Astrophysical Observatory of the Russian Academy of Sciences, the 8.2-m VLT telescope of the European Southern Observatory at Paranal (Chile), the IRAM 30-m and Plateau de Bure Telescopes and the 100-m telescope of the Max-Planck-Institut für Radioastronomie at Effelsberg (Germany). IRAM is supported by INSU/CNRS (France), MPG (Germany) and IGN (Spain). We thank both the SAO-RAS Director and the ESO Director's Discretionary Time Committee for accepting the observation a few days after the onset of the event. We also thank N. Schartel for allotting XMM-Newton DDT time for a late-time X-ray observation. We acknowledge the use of public

data from the Swift data archives and the service provided by the γ-ray burst Coordinates Network (GRB) and BACODINE system, maintained by S. Barthelmy. A.J.C.T. acknowledges discussions with E. Alfaro, Y. Beletski, A. M. Belobodorov, W. Cui, S. Digel, R. Fernández-Muñoz, P. Gandhi, S. Gottlieb, M. Lyutikov, A. Merloni, M. A. Pérez-Torres, V. Reglero, E. Ros and G. Sala. P.F., D.A.K. and S.K. acknowledge financial support by DFG and D.P.R. from Junta de Andalucía. L.H. acknowledges support from IRCSET and SFI. R.H. acknowledges support from PECs and GACR. This work was partially supported by the Spanish Ministry of Science and Innovation.

Author Contributions A.J.C.T. wrote the paper. A.U.P., D.A.K., S. S. and A.C.W. prepared the figures and together with M. B., P. F., A. K., S. K., S. L. and D.P.R. participated in the data analysis. S. M. and K. L. helped in the modelling. The remaining coauthors provided observational data, discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.J.C.T. (ajct@iaa.es).

Vol 455|25 September 2008|doi:10.1038/nature07288

LETTERS

Reconstruction of non-classical cavity field states with snapshots of their decoherence

Samuel Deléglise[1], Igor Dotsenko[1,2], Clément Sayrin[1], Julien Bernu[1], Michel Brune[1], Jean-Michel Raimond[1] & Serge Haroche[1,2]

The state of a microscopic system encodes its complete quantum description, from which the probabilities of all measurement outcomes are inferred. Being a statistical concept, the state cannot be obtained from a single system realization, but can instead be reconstructed[1] from an ensemble of copies through measurements on different realizations[2–4]. Reconstructing the state of a set of trapped particles shielded from their environment is an important step in the investigation of the quantum–classical boundary[5]. Although trapped-atom state reconstructions[6–8] have been achieved, it is challenging to perform similar experiments with trapped photons because cavities that can store light for very long times are required. Here we report the complete reconstruction and pictorial representation of a variety of radiation states trapped in a cavity in which several photons survive long enough to be repeatedly measured. Atoms crossing the cavity one by one are used to extract information about the field. We obtain images of coherent states[9], Fock states with a definite photon number and 'Schrödinger cat' states (superpositions of coherent states with different phases[10]). These states are equivalently represented by their density matrices or Wigner functions[11]. Quasi-classical coherent states have a Gaussian-shaped Wigner function, whereas the Wigner functions of Fock and Schrödinger cat states show oscillations and negativities revealing quantum interferences. Cavity damping induces decoherence that quickly washes out such oscillations[5]. We observe this process and follow the evolution of decoherence by reconstructing snapshots of Schrödinger cat states at successive times. Our reconstruction procedure is a useful tool for further decoherence and quantum feedback studies of fields trapped in one or two cavities.

Engineering and reconstructing non-classical states of trapped light requires cavities that prevent the escape of even single photons during the preparation and read-out procedures. We have built a cavity made of highly reflecting superconducting mirrors[12] whose long damping time, $T_c = 0.13$ s, allows the trapped field to interact with thousands of atoms crossing the cavity one by one. The interaction with atoms is used to turn an initial coherent field into a Fock or Schrödinger cat state and, subsequently, to reconstruct it. An ensemble of trapped photons becomes, like a collection of trapped atoms, an 'object of investigation' to be manipulated and observed for fundamental tests and quantum information purposes.

Our set-up is sketched in Fig. 1a. The cavity C, resonant at 51 GHz, is cooled to a temperature of 0.8 K (mean number of residual blackbody photons, $n_b = 0.05$). A coherent microwave field with a Poisson photon number distribution (mean, n_m; standard deviation, $\Delta n = \sqrt{n_m}$) is initially injected into C using a classical pulsed source S. Rubidium atoms from an atomic beam are prepared, in box B, in the circular Rydberg state with principal quantum number 50 ($|g\rangle$). The cavity is detuned from the transition between $|g\rangle$ and the adjacent circular state with principal quantum number 51 ($|e\rangle$) by an amount δ, precluding atom–field energy exchange. The pulsed atom preparation produces Rydberg atoms with a velocity of 250 m s^{-1}. Auxiliary microwave cavities R_1 and R_2 sandwiching C are connected to a microwave source S′. They are used to apply resonant pulses to the atoms. The R_1 pulse performs the $|g\rangle \rightarrow (|e\rangle + |g\rangle)/\sqrt{2}$ transformation. The same pulse, differing by an adjustable phase shift ϕ, is applied in R_2. Atoms are counted by the detector D, discriminating $|e\rangle$ and $|g\rangle$ (one atom on average every 0.5 ms). For experimental details, see refs 13 and 14.

The R_1–R_2 combination forms a Ramsey interferometer[14]. It is sensitive to the atomic-state-superposition phase shift induced by the atom's interaction with the field in C, which is characterized by

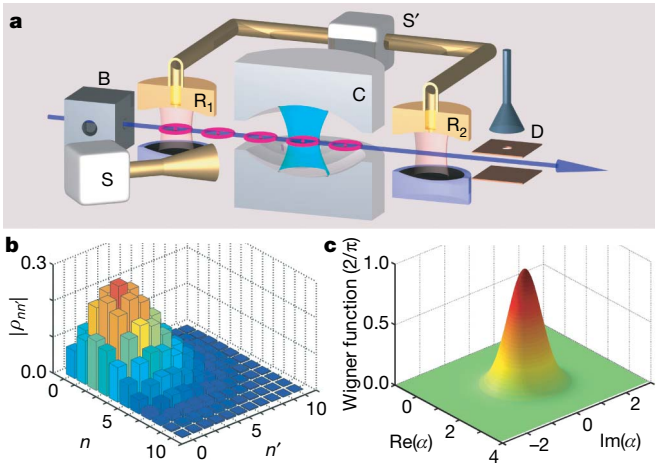


Figure 1 | Reconstructing a coherent state. a, The set-up, showing the stream of atoms prepared in box B and crossing the R_1–R_2 interferometer in which the cavity C, made of two mirrors facing each other, is inserted. The source S, coupled to a waveguide, generates a coherent microwave pulse irradiating C on one side. By diffraction at the mirrors' edges, it injects into C a small coherent field with controlled amplitude and phase. The field outside vanishes almost instantaneously when S is switched off. The source S is used to prepare an initial field in C and, later, to translate the field for state reconstruction. Another pulsed source, S′, feeds the interferometer cavities R_1 and R_2. Information is extracted from the field by state-selective atomic counting in D. **b**, Density matrix (modulus of matrix elements) of a coherent state of amplitude $\beta = \sqrt{2.5}$, reconstructed in an 11-dimensional Hilbert space. The reconstruction parameters are $\delta/2\pi = 65$ kHz and $\phi = -\Phi(0, \delta) + \pi$. We sample 161 points in phase space and for each point detect ~7,000 atoms over 600 realizations. The fidelity, $F = \langle\beta|\rho|\beta\rangle$, of the reconstructed state is 0.98. **c**, Wigner function (in units of $2/\pi$) obtained from the density matrix shown in **b**.

[1]Laboratoire Kastler Brossel, Ecole Normale Supérieure, CNRS, Université Pierre et Marie Curie, 24 rue Lhomond, 75231 Paris Cedex 05, France. [2]Collège de France, 11 place Marcelin Berthelot, 75231 Paris Cedex 05, France.

the Rabi frequency $\Omega/2\pi = 49$ kHz. This phase shift is described by an operator $\Phi(N,\delta)$ that depends on δ and the photon number operator $N = a^\dagger a$ (where a and $a^\dagger$ respectively denote the photon annihilation and creation operators). To lowest order, $\Phi(N,\delta)$ is linear in N, but for the small δ/Ω values of our experiment, we take into account its exact nonlinear expression[13]. The interferometer measures $\cos(\Phi(N,\delta)+\phi)$, which is sensitive to the diagonal elements of the field density matrix in the Fock state basis, but contains no information about the coherences between these states. To obtain this information, we measure the phase shifts produced by the field after it has been translated in phase space, by mixing it with reference coherent fields of adjustable complex amplitudes α. These translations, described by the operators $D(\alpha) = \exp(\alpha a^\dagger - \alpha^* a)$ (where the asterisk denotes complex conjugation), are achieved by injecting a second field pulse into C.

We denote by ρ the density matrix of the field to be reconstructed (Fock-state-basis matrix elements, $\rho_{nn'}$), by $\rho^{(\alpha)} = D(\alpha)\rho D(-\alpha)$ the density matrix after field translation and by P_e and P_g the respective probabilities of finding in $|e\rangle$ and $|g\rangle$ the first atom that crosses the interferometer (experimentally obtained by averaging over many field realizations). The difference $P_e - P_g = \mathrm{Tr}[\rho^{(\alpha)}\cos(\Phi(N,\delta)+\phi)]$ is the expectation value in the translated state of the diagonal field operator $\cos(\Phi(N,\delta)+\phi)$. The measurement is non-demolition for the photon number[15], and the ensemble average of first-crossing atoms does not change $\rho^{(\alpha)}_{nn}$. Hence, the same $P_e - P_g$ expression holds for the second (or any subsequent) atom. We thus determine $P_e - P_g$ by averaging the detections of successive atoms for a single field realization with those measured for different realizations. A measuring sequence for each realization lasts 4 ms, which is short in comparison with the characteristic evolution time of the state. We also correct the raw $P_e - P_g$ values by taking into account the known imperfections of the interferometer.

The $P_e - P_g$ difference is also the expectation value of $G(\alpha,\phi,\delta) = D(-\alpha)\cos(\Phi(N,\delta)+\phi)D(\alpha)$ in state ρ. By sampling α values, we obtain the expectations $g(\alpha,\phi,\delta)$ of an ensemble of non-commuting $G(\alpha,\phi,\delta)$ operators satisfying:

$$\mathrm{Tr}[\rho G(\alpha,\phi,\delta)] = g(\alpha,\phi,\delta) \quad (1)$$

Provided that we sample a large enough number of points α in phase space, equation (1) allows us to reconstruct ρ. To ensure that the reconstructed state does not contain any information other than that extracted from the data, we also maximize the field entropy, $-\mathrm{Tr}[\rho\ln\rho]$, during the reconstruction procedure (principle of maximum entropy[16]).

The Wigner function associated with the state described by the density matrix ρ is defined at point α in phase space as[14]

$$W(\alpha) = \frac{2}{\pi}\mathrm{Tr}[D(-\alpha)\rho D(\alpha)e^{i\pi N}] \quad (2)$$

and (to within a normalization) is the expectation value of the photon number parity operator $\exp(i\pi N)$ in the state translated by $-\alpha$. The Wigner function could be determined directly[17] if the atoms underwent an exact phase shift of π per photon, realizing the measurement of $\exp(i\pi N)$ after field translation by different values of α. This would be a special case of reconstruction corresponding to $\Phi(N,\delta) - \Phi(0,\delta) = \pi N$. This relation cannot be satisfied, owing to nonlinear atom–field phase shift. Instead of a direct determination of the Wigner function, we thus reconstruct ρ first and then obtain the Wigner function using equation (2).

Figure 1b shows the reconstructed density matrix of a coherent state. Along its diagonal, we recognize the Poisson photon number distribution ρ_{nn}. The off-diagonal elements describe the classical coherence of the state. The corresponding Wigner function, shown in Fig. 1c, is, as expected, a Gaussian peak with a circular symmetry.

As a first non-classical example, we have reconstructed Fock states. To generate them, we prepare a coherent field and let it interact with atoms, achieving a quantum non-demolition measurement of the photon number that progressively projects the field onto a Fock state $|n_0\rangle$. This procedure is adapted from ref. 15, taking into account the known effect of cavity damping during state projection. Following this preparation, we apply our reconstruction method with subsequent probe atoms and reconstruct the Fock states present in the expansion of the initial coherent state.

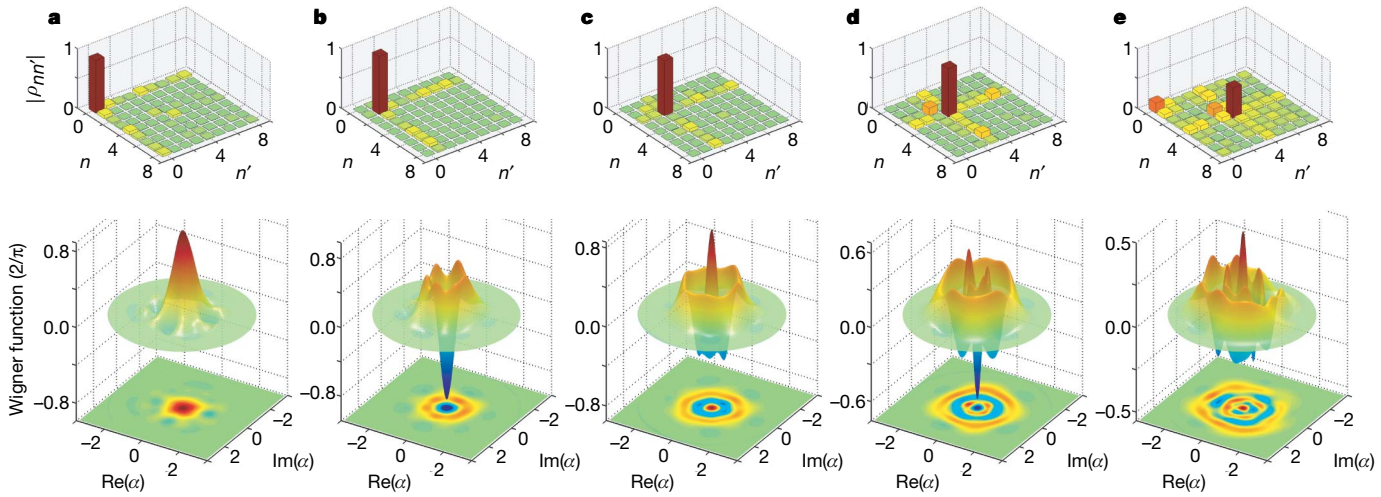


Figure 2 | Reconstructing Fock states. a–e, Reconstructed density matrices (modulus of matrix elements) and Wigner functions (in units of $2/\pi$) of the $n_0 = 0, 1, 2, 3, 4$ Fock states prepared by quantum non-demolition projection of an initial coherent field ($n_m = 1.5$ for $n_0 = 0, 1, 2, 3$; $n_m = 5.5$ for $n_0 = 4$). The Wigner functions are shown as three-dimensional plots and two-dimensional projections. We select a photon number n_0 if, after the detection of ~60 preparation atoms, the measurement has converged to a Fock state having a probability >0.9 for $n_0 = 0, 1, 2, 3$ and >0.8 for $n_0 = 4$. The same detuning ($\delta/2\pi = 120$ kHz) is used for state preparation and reconstruction, corresponding to $\partial\Phi/\partial n \approx \pi/2$ at $n = 3$. Two values of ϕ ($-\Phi(0,\delta)+\pi$ and $-\Phi(0,\delta)+\pi/2$) are used for state preparation and reconstruction, which is made in a nine-dimensional Hilbert space. We sample ~400 points in phase space and, for each point, average between about 50 (for $n_0 = 3$) and 500 ($n_0 = 0, 1$) realizations, with ~10 atoms in each realization. In addition to the main peak in the density matrices, for $n_0 > 1$ we observe a small diagonal peak at $n = n_0 - 1$ due to cavity damping during reconstruction. A peak at $n = 0$ also appears in the $n_0 = 4$ density matrix, because of imperfections in the state preparation process that selects the photon number modulo four (as $\Phi(n+4,\delta) \approx \Phi(n,\delta) + 2\pi$). The off-diagonal elements in the density matrices and the corresponding fluctuations in the angular distributions of the Wigner functions mainly reflect statistical noise (fewer atoms detected for reconstructing $n_0 = 2, 3, 4$ than for $n_0 = 0, 1$). The fidelities, $F = \langle n_0|\rho|n_0\rangle$, of the reconstructed states are 0.89, 0.98, 0.92, 0.82, 0.51 for $n_0 = 0, 1, 2, 3, 4$, respectively.

Figure 2 displays the obtained density matrices together with the corresponding Wigner functions for $n_0 = 0$ (vacuum), 1, 2, 3, 4. As expected, the density matrices mainly exhibit a single diagonal peak. Each Wigner function shows circular rings around the origin of phase space, where it is positive for even n_0 and negative for odd n_0. The number of rings and their size increases as expected with n_0. Photonic Fock states with small values of n_0 have already been reconstructed in free space[18–20] and in a cavity[21], but to our knowledge this is the first Fock state reconstruction with $n_0 > 2$.

To generate a Schrödinger cat state[22], we first inject into C a coherent field of amplitude $\beta = \sqrt{n_m}$. We then prepare an atom in the state $(|e\rangle + |g\rangle)/\sqrt{2}$ using R_1 and send it into C. The two atomic components shift the phase of the field in opposite directions. Neglecting atom–field phase shift nonlinearity, the field is split into two coherent states of complex amplitudes $\beta\exp(\pm i\chi)$, where $\chi = (\partial\Phi(n,\delta)/\partial n)/2$ evaluated at $n = n_m$. The atom is entangled with the field in the state $(|e\rangle|\beta\exp(i\chi)\rangle + |g\rangle|\beta\exp(-i\chi)\rangle)/\sqrt{2}$. The R_2 pulse then mixes $|e\rangle$ and $|g\rangle$ again. Last, the atomic detection, depending upon its outcome ($|g\rangle$ or $|e\rangle$), projects the field onto one of the two Schrödinger cat states $(|\beta\exp(i\chi)\rangle \pm |\beta\exp(-i\chi)\rangle)/\sqrt{2}$. We call these 'even' (plus sign) and 'odd' (minus sign) Schrödinger cat states because they contain, for $\chi = \pi/2$, only even and odd photon numbers, respectively. After this preparation, we apply our reconstruction procedure.

Figure 3a, b shows the Wigner functions of the even and odd Schrödinger cat states obtained from the same coherent field ($n_m = 3.5$, $\chi = 0.37\pi$). They exhibit two well-separated positive peaks that are associated with the classical components and whose slightly elongated shapes are due to the phase shift nonlinearity neglected above. The 'size' of each of these states, defined as the squared distance between the peaks (and expressed as a number of photons), is $d^2 \approx 4n_m\sin^2\chi = 11.8$ photons. Between these peaks, oscillating features with alternating positive and negative values are the signatures of the state's quantum interference. The even and odd Schrödinger cat states have nearly identical classical components and only differ in the sign of their quantum interferences. The theoretical Wigner functions, taking the nonlinearity in the preparation of the Schrodinger cat states into account, are shown for comparison in the insets. The fidelity of the two states (overlap between the reconstructed ρ and the expected ρ) is $F = 0.72$. It is mainly limited by imperfections in the R_1 and R_2 pulses applied to the preparation atom, which reduce the contrast of the quantum interference feature. If the preparing atom is detected without discriminating $|e\rangle$ and $|g\rangle$, we obtain the statistical mixture of even and odd Schrödinger cat states whose Wigner function is shown in Fig. 3c. Equivalently, this is a statistical mixture of the two classical components. Although non-classical states of propagating light with similar Wigner functions have previously been observed[23], here well-separated classical components of a field can be identified in a reconstructed state and unambiguously distinguished from their quantum interference term.

Schrödinger cats states are model states for exploring decoherence, the phenomenon accounting for the transition between quantum and classical behaviours[5]. Our reconstruction method allows us to study this process. Immediately after state preparation, we realize the $D(\alpha)$ translation and detect a sequence of atoms divided into 4-ms time windows. These atoms record $P_e - P_g$ as a function of time, without modifying the dynamics of this quantity. We average the results of realizations corresponding to the same translation and time window, and then repeat the process for different values of α. This provides a direct record of the evolution of the translated states, rather than that of the state itself. The two dynamics are closely related, however. Decoherence acting on the initial density operator $\rho(0)$ turns it at time t into $\rho(t) = L[\rho(0), t]$. Here L is the decoherence superoperator[14], which can be shown to satisfy the following relation:

$$D(\alpha e^{-t/2T_c})L[\rho(0),t]D(-\alpha e^{-t/2T_c}) = L[D(\alpha)\rho(0)D(-\alpha),t]$$

Translating the initial field by α and letting it evolve over time t is equivalent to letting it evolve during that time and translating it by $\alpha\exp(-t/2T_c)$. We thus analyse the data obtained at time t as if they corresponded to a translation rescaled by $\exp(-t/2T_c)$. This is more efficient than letting the field evolve before translating it, because we exploit all the data of a long sequence instead of recording only a short time window for each delay. We have experimentally checked the equivalence between the two methods by comparing the results for one time delay and have verified that the reconstructed Schrödinger cat states are, within the effects of noise, undistinguishable.

Figure 4a shows four snapshots of the Wigner function of an odd Schrödinger cat at increasing times, clearly revealing decoherence. Although the classical components have hardly decayed after 50 ms,

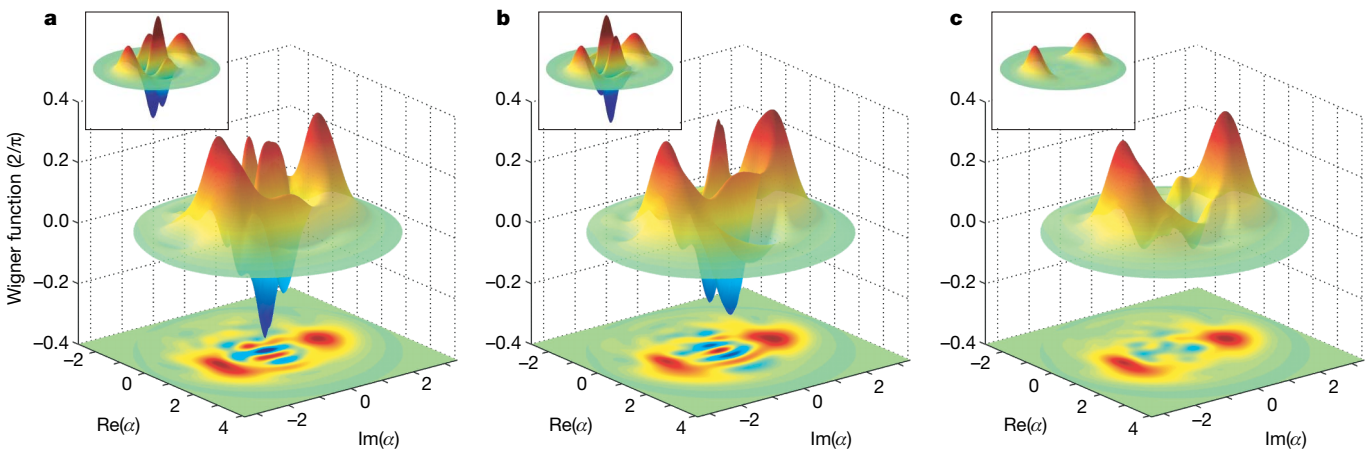


Figure 3 | Reconstructing Schrödinger cat states. a, b, The Wigner functions (in units of $2/\pi$) of even (**a**) and odd (**b**) Schrödinger cat states with $n_m = 3.5$ and $\chi = 0.37\pi$ are reconstructed, following state preparation. The same detuning ($\delta/2\pi = 51$ kHz) and interferometer phase ($\phi = -\Phi(0,\delta) + \pi$) are used for state preparation and reconstruction. The number of sampling points is ~500, with ~2,000 atoms detected at each point, in 400 realizations. The dimension of the Hilbert space used for reconstruction is 11. The small insets present for comparison the theoretical Wigner functions computed in the case of ideal preparation and detection of the atomic state superpositions. Decoherence during state preparation is taken into account. The maximum theoretical values of the classical components and interference fringes are close to 0.5 and 1, respectively. In the reconstructed states, the quantum interference is smaller, mainly owing to imperfections of the Ramsey interferometer that affect the preparation of the Schrödinger cat state (but not its reconstruction). **c**, Reconstructed Wigner function of the field prepared in C when the state of the preparation atom is not read-out (statistical mixture of two classical fields). Inset, corresponding theoretical Wigner function.

the interference feature has vanished, turning the initial state into a statistical mixture similar to that shown in Fig. 3c. A complete movie of the evolution of the Wigner function of a Schrödinger cat state is presented in the Supplementary Information. By subtracting the Wigner functions of the even and odd Schrödinger cat states corresponding to the same preparation sequence, we isolate their interference features by cancelling their equal, classical, parts. A movie showing the progressive vanishing of this interference is also provided in the Supplementary Information.

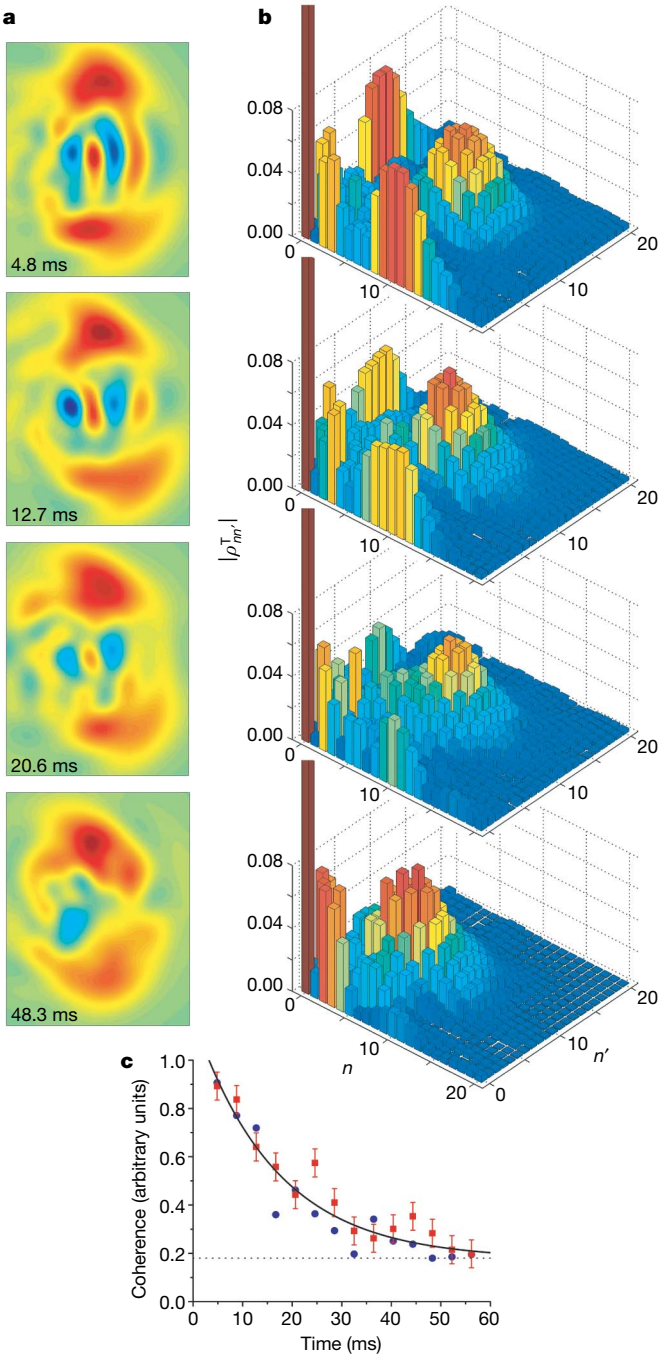


Figure 4 | Movie of decoherence. a, Snapshots of the Wigner function of the odd Schrödinger cat state in Fig. 3b at four successive times after state preparation. **b**, Corresponding snapshots of the translated density matrix ρ^{T} (modulus of matrix elements). The dimension of the Hilbert space is enlarged to 30. The $n = 0$ peak is clipped, its amplitude being ~0.38 in each snapshot. In all frames, the small matrix elements for $n < 5$ are due to the deviation of the classical component of the Schrödinger cat state from an ideal coherent state. **c**, The quantum coherence of the Schrödinger cat states (even, red; odd, blue), defined as the sum of the $|\rho^{T}_{n0}|$ for $n \geq 5$, plotted as a function of time. The statistical error bars (shown, for clarity, for the even Schrödinger cat states only) are obtained by analysing the state reconstructions performed on different subsamples of measured data. The solid line is a common exponential fit, including an offset (dotted line) accounting for residual noise in the modulus of the density matrix elements.

It is also instructive to observe decoherence directly in the density matrix. To distinguish the classical coherence of each component of the Schrödinger cat state from their mutual quantum coherence, we consider the mathematically translated reconstructed density matrix $\rho^{T} = \rho^{-\beta\exp(i\chi)}$, whose classical components are close to the vacuum $|0\rangle$ and to $|-2i\beta\sin\chi\rangle$. This formal translation leaves unchanged the distance between the two classical components in the phase plane as well as their mutual coherence.

In Fig. 4b, we present the density matrix $\rho^{T}(t)$ of the Schrödinger cat state in Fig. 4a, reconstructed for the same times. In each frame, the diagonal elements present two maxima around $n = 0$ and $n = 11$. The off-diagonal elements are of two kinds. Those for which $|n - n'| \approx \sqrt{11}$ describe the classical coherence of the non-vacuum component and remain nearly unchanged on the observed timescale. The off-diagonal terms in the first row and column of the matrix (respectively ρ^{T}_{0n} and ρ^{T}_{n0}) initially exhibit a bell-shaped variation with n, centred at $n \approx 11$. These terms correspond to the quantum coherence of the Schrödinger cat state responsible for the oscillations observed in the Wigner function. Their fast decay is the signature of decoherence.

The measured quantum coherence of the even and odd Schrödinger cat states is plotted as a function of time in Fig. 4c. A common exponential fit yields a decoherence time of $T_d = 17 \pm 3$ ms. A simple analytical model of decoherence[14] predicts that $T_d = 2T_c/d^2 = 22$ ms at $T = 0$ K, reduced to[24] $T_d = 2T_c/(d^2(1+2n_b)+4n_b) = 19.5$ ms when including a thermal background at $T = 0.8$ K, in good agreement with the measured value. A movie of a smaller Schrödinger cat state ($d^2 = 8$) yields $T_d = 28$ ms, illustrating the dependence of the decoherence time on the Schrödinger cat state size[5,14]. Earlier experiments have studied the relaxation of photonic[22] and atomic[25] Schrödinger cat states by observing some of their specific features, but this experiment allows us to produce a movie of decoherence in a fully reconstructed Schrödinger cat state.

We have shown that atoms interacting with a cavity field can be used to engineer and reconstruct a wide variety of photonic states and to study their evolution. Going one step farther, we plan to use information provided by the atoms to implement feedback procedures and preserve the quantum coherence over longer time intervals[26]. We will also extend these studies to fields stored in two cavities. Atoms will be used to entangle the cavity fields into non-local quantum states[27,28], reconstruct these states and protect them against decoherence by means of quantum feedback operations.

Received 25 June; accepted 23 July 2008.

1. Paris, M. G. A. & Řeháček, J. (eds) *Quantum State Estimation* (Springer, 2004).
2. Smithey, D. T., Beck, M., Raymer, M. G. & Faridani, A. Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum. *Phys. Rev. Lett.* **70**, 1244–1247 (1993).
3. Dunn, T. J., Walmsley, I. A. & Mukamel, S. Experimental determination of the quantum-mechanical state of a molecular vibrational mode using fluorescence tomography. *Phys. Rev. Lett.* **74**, 884–887 (1995).
4. Kurtsiefer, C., Pfau, T. & Mlynek, J. Measurement of the Wigner function of an ensemble of helium atoms. *Nature* **386**, 150–153 (1997).
5. Zurek, W. Decoherence, einselection, and the quantum origins of the classical. *Rev. Mod. Phys.* **75**, 715–775 (2003).
6. Leibfried, D. *et al.* Experimental determination of the motional quantum state of a trapped atom. *Phys. Rev. Lett.* **77**, 4281–4285 (1996).
7. Häffner, H. *et al.* Scalable multiparticle entanglement of trapped ions. *Nature* **438**, 643–646 (2005).
8. Morinaga, M., Bouchoule, I., Karam, J.-C. & Salomon, C. Manipulation of motional quantum states of neutral atoms. *Phys. Rev. Lett.* **83**, 4037–4040 (1999).
9. Glauber, R. J. Coherent and incoherent states of the radiation field. *Phys. Rev.* **131**, 2766–2788 (1963).

10. Bužek, V. & Knight, P. L. Quantum interference, superposition states of light, and nonclassical effects, in *Progress in Optics XXXIV* (ed. Wolf, E.) 1–158 (Elsevier, 1995).
11. Schleich, W. P. *Quantum Optics in Phase Space* (Wiley, 2001).
12. Kuhr, S. *et al.* Ultrahigh finesse Fabry-Pérot superconducting resonator. *Appl. Phys. Lett.* **90,** 164101 (2007).
13. Raimond, J.-M., Brune, M. & Haroche, S. Colloquium: manipulating quantum entanglement with atoms and photons in a cavity. *Rev. Mod. Phys.* **73,** 565–582 (2001).
14. Haroche, S. & Raimond, J.-M. *Exploring the Quantum: Atoms, Cavities and Photons* (Oxford Univ. Press, 2006).
15. Guerlin, C. *et al.* Progressive field state collapse and quantum non-demolition photon counting. *Nature* **448,** 889–893 (2007).
16. Bužek, V. & Drobny, G. Quantum tomography via the *MaxEnt* principle. *J. Mod. Opt.* **47,** 2823–2839 (2000).
17. Lutterbach, L. G. & Davidovich, L. Method for direct measurement of the Wigner function in cavity QED and ion traps. *Phys. Rev. Lett.* **78,** 2547–2550 (1997).
18. Lvovsky, A. I. *et al.* Quantum state reconstruction of the single-photon Fock state. *Phys. Rev. Lett.* **87,** 050402 (2001).
19. Zavatta, A., Viciani, S. & Bellini, M. Tomographic reconstruction of the single-photon Fock state by high-frequency homodyne detection. *Phys. Rev. A* **70,** 053821 (2004).
20. Ourjoumtsev, A., Tualle-Brouri, R. & Grangier, P. Quantum homodyne tomography of a two-photon Fock state. *Phys. Rev. Lett.* **96,** 213601 (2006).
21. Bertet, P. *et al.* Direct measurement of the Wigner function of a one-photon Fock state in a cavity. *Phys. Rev. Lett.* **89,** 200402 (2002).
22. Brune, M. *et al.* Observing the progressive decoherence of the 'meter' in a quantum measurement. *Phys. Rev. Lett.* **77,** 4887–4890 (1996).
23. Ourjoumtsev, A., Jeong, H., Tualle-Brouri, R. & Grangier, P. Generation of optical 'Schrödinger cats' from photon number states. *Nature* **448,** 784–786 (2007).
24. Kim, M. S. & Bužek, V. Schrödinger cat states at finite temperature: Influence of a finite temperature heat bath on quantum interferences. *Phys. Rev. A.* **46,** 4239–4251 (1992).
25. Myatt, C. J. *et al.* Decoherence of quantum superpositions through coupling to engineered reservoirs. *Nature* **403,** 269–273 (2000).
26. Zippilli, S., Vitali, D., Tombesi, P. & Raimond, J.-M. Scheme for decoherence control in microwave cavities. *Phys. Rev. A* **67,** 052101 (2003).
27. Davidovich, L., Brune, M., Raimond, J.-M. & Haroche, S. Mesoscopic quantum coherences in cavity QED: preparation and decoherence monitoring schemes. *Phys. Rev. A* **53,** 1295–1309 (1996).
28. Milman, P. *et al.* A proposal to test Bell's inequalities with mesoscopic non-local states in cavity QED. *Eur. Phys. J. D* **32,** 233239 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by the Agence Nationale pour la Recherche (ANR), by the Japan Science and Technology Agency (JST) and by the European Union under the Integrated Projects SCALA and CONQUEST. S.D. is funded by the Délégation Générale pour l'Armement (DGA).

Author Contributions S.D., I.D. and C.S. contributed equally to this work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to S.H. (haroche@lkb.ens.fr).

LETTERS

Magnetization vector manipulation by electric fields

D. Chiba[1,2], M. Sawicki[2,3], Y. Nishitani[2], Y. Nakatani[4], F. Matsukura[2,1] & H. Ohno[2,1]

Conventional semiconductor devices use electric fields to control conductivity, a scalar quantity, for information processing. In magnetic materials, the direction of magnetization, a vector quantity, is of fundamental importance. In magnetic data storage, magnetization is manipulated with a current-generated magnetic field (Oersted–Ampère field), and spin current[1,2] is being studied for use in non-volatile magnetic memories[3,4]. To make control of magnetization fully compatible with semiconductor devices, it is highly desirable to control magnetization using electric fields. Conventionally, this is achieved by means of magnetostriction produced by mechanically generated strain through the use of piezoelectricity[5–8]. Multiferroics[9,10] have been widely studied in an alternative approach where ferroelectricity is combined with ferromagnetism. Magnetic-field control of electric polarization has been reported in these multiferroics using the magnetoelectric effect, but the inverse effect—direct electrical control of magnetization—has not so far been observed[11]. Here we show that the manipulation of magnetization can be achieved solely by electric fields in a ferromagnetic semiconductor, (Ga,Mn)As. The magnetic anisotropy, which determines the magnetization direction, depends on the charge carrier (hole) concentration in (Ga,Mn)As. By applying an electric field using a metal–insulator–semiconductor structure[12–14], the hole concentration and, thereby, the magnetic anisotropy can be controlled, allowing manipulation of the magnetization direction.

Our device structure (Fig. 1a) consists of a thin ferromagnetic (Ga,Mn)As layer grown on a GaAs (001) substrate (substrate not shown in Fig. 1a; see Methods Summary for the detailed structure) with an insulator and a metal gate electrode on top so that an electric field, **E**, can be applied to the (Ga,Mn)As layer. The magnetic easy axis, along which the magnetic anisotropy energy is the lowest when the magnetization, **M**, is collinear with the axis, lies in the plane owing to compressive strain in the film[15]. The magnetic easy axis is dependent on the hole concentration, p, and can be modulated by the application of **E**, changing p, which thus allows manipulation of **M** purely through the use of an electric field (Fig. 1a). In the present case, the situation is made slightly more complex by the presence of two major magnetic easy axes within the plane. The competition between the two magnetic easy axes, namely biaxial easy axes (along the [100] and [010] directions) and a uniaxial easy axis (along [110] or [$\bar{1}$10]) (refs 16–22; see Fig. 1b), determines the actual direction of **M**.

Transport measurements have been used to detect the direction of **M** under the application of **E**. Figure 1c shows our device for transport measurement and the measurement configuration. Two cleaved, as-grown samples (devices A and B) were processed into a Hall bar geometry along [110] with a channel of 40-μm width and

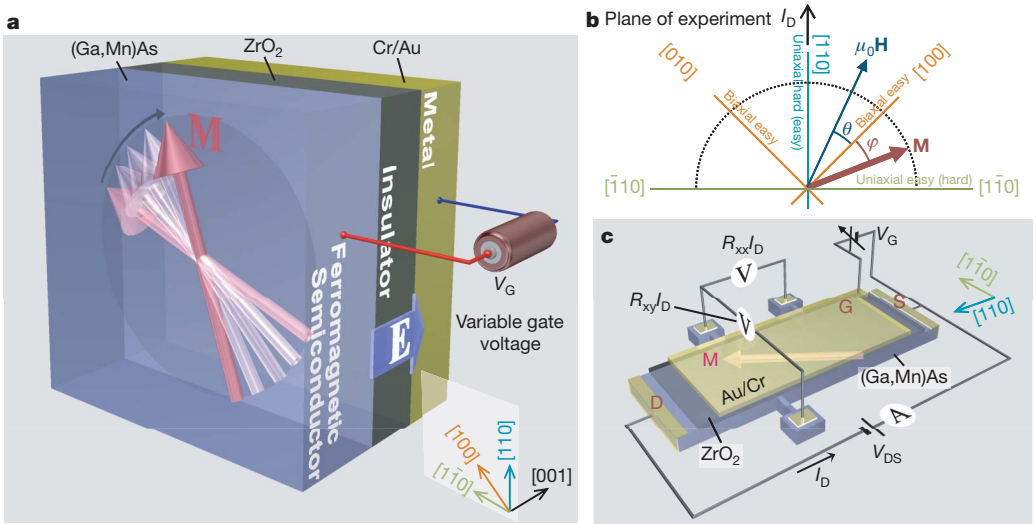


Figure 1 | Magnetization vector manipulation by electric field, and measurement configuration. a, The device comprises a metal gate (Cr/Au), an insulator layer (ZrO_2), and a ferromagnetic semiconductor (Ga,Mn)As layer. The gate voltage is applied between the metal gate and the (Ga,Mn)As layer and induces the electric field (**E**), which results in the change of the hole concentration, p, in the ferromagnetic semiconductor. The orientation of the magnetization (**M**) varies as a consequence of anisotropy modulation by p. **b**, Vector diagram of **M** and the external magnetic field, μ_0**H**. The angles φ and θ are the angles between **M** and [100] and μ_0**H** and [100], respectively. I_D, drain current. **c**, The Hall-bar-shaped device. The Hall bar is used to measure longitudinal and transverse resistance (respectively R_{xx} and R_{xy}) under the variable gate voltage V_G and μ_0**H**. S, source; D, drain; G, gate. A constant S–D voltage (V_{DS} = 1 V) is applied during measurement.

[1]Semiconductor Spintronics Project, Exploratory Research for Advanced Technology, Japan Science and Technology Agency, Sanban-cho 5, Chiyoda-ku, Tokyo 102-0075, Japan. [2]Laboratory for Nanoelectronics and Spintronics, Research Institute of Electrical Communication, Tohoku University, Katahira 2-1-1, Aoba-ku, Sendai 980-8577, Japan. [3]Institute of Physics, Polish Academy of Sciences, Al. Lotników 32/46, PL-02668, Warszawa, Poland. [4]University of Electro-communications, Chofugaoka 1-5-1, Chofu, Tokyo 182-8585, Japan.

140-μm length. 5-nm Cr/100-nm Au electrode pads were evaporated and lifted off to form voltage probes and source (S) and drain (D) contacts. Subsequently, the samples were placed in an atomic-layer-deposition chamber and a 31-nm ZrO_2 layer was deposited as a gate insulator at a substrate temperature of 120 °C for device A and a 41-nm ZrO_2 layer was deposited as a gate insulator at a substrate temperature of 150 °C for device B. Lastly, a 5-nm Cr/100-nm Au gate (G) electrode was evaporated and lifted off. The dielectric constant of the ZrO_2 film deposited at 150 °C was determined to be 23 from separate measurements. Four-terminal measurements were performed by applying a constant, d.c. D–S voltage $V_{DS} = 1$ V to probe the longitudinal (sheet, R_{xx}) and transverse (planar-Hall or Hall, R_{xy}) components of the resistance as functions of external magnetic field $\mu_0\mathbf{H}$ (where μ_0 is the permeability of vacuum) and gate voltage V_G (Fig. 1c). The maximum value of $|V_G|$ was 12 V, which corresponds to $|\mathbf{E}|$ values of 3.9 MV cm^{-1} for device A and 2.9 MV cm^{-1} for device B.

We first focus on the in-plane rotation of **M** in device A. An in-plane $\mu_0\mathbf{H}$ was rotated around the device and the dependence of R_{xy} on θ, the angle between $\mu_0\mathbf{H}$ and [100], was measured. From this we determined φ, the angle between **M** and [100], in the following way (θ and φ are shown in Fig. 1b). R_{xy} has been experimentally shown to depend on φ (planar Hall effect; see Supplementary Information):

$$R_{xy} = R_{xy0}\cos 2\varphi$$

Here R_{xy0} is the amplitude of R_{xy}. Let E denote the out-of-plane component of **E** and let $H = |\mathbf{H}|$. The thick solid line in Fig. 2a represents R_{xy}/R_{xy0} measured as a function of θ for $E = -3.9$ MV cm^{-1} and $\mu_0 H = 0.15$ T at 2 K, taken by sweeping θ anticlockwise (the direction in which θ increases). The shape of the obtained curve is distorted from the cos2θ curve (Fig. 2a, dotted line), showing that φ differs from θ. This difference is due to the presence of magnetic anisotropy. When **M** lies along the easy axis, the magnetic anisotropy energy is minimized (as discussed above), making the direction of **M** insensitive to the direction of the external (weak) magnetic field. Expressing φ as a function of θ (Fig. 2b, thick solid line) shows that φ remains θ closer to 0° ([100]) or 90° ([010]) as θ is increased, showing that these two directions are biaxial easy axes. Conversely,

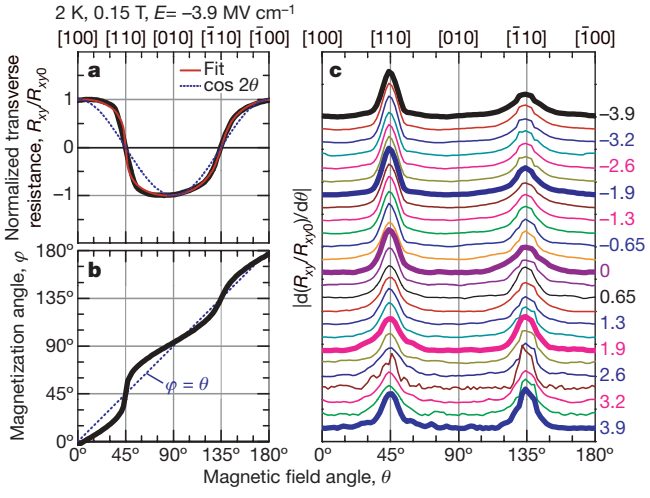


Figure 2 | Dependence of transverse resistance and magnetization angle on magnetic field angle. a, The dependence of the normalized planar Hall resistance R_{xy}/R_{xy0} on the magnetic field angle, θ, for $E = -3.9$ MV cm^{-1}, at 2 K (thick solid line). A constant magnetic field of $\mu_0 H = 0.15$ T is applied. The thin solid line shows the fitting curve, from which the in-plane anisotropy fields are determined. The dotted line indicates R_{xy}/R_{xy0} when magnetization angle, φ, equals θ. **b**, φ as a function of θ (thick solid line). The dotted line is the case $\varphi = \theta$. **c**, The absolute value of the derivative of R_{xy}/R_{xy0} with respect to θ for values of E from -3.9 MV cm^{-1} to 3.9 MV cm^{-1} (labelling the lines in units of MV cm^{-1}), at 2 K and 0.15 T. Two peaks are observed and the higher one is the harder axis for **M**.

the curve intersects the line $\varphi = \theta$ (Fig. 2b, dotted line) steeply at $\theta = 45°$ ([110]) and 135° ([$\bar{1}$ 10]), indicating the hard axes.

Figure 2c shows the absolute value of the derivative with respect to θ of the normalized R_{xy} curves, for various values of E. We observe two peaks of different heights at $\theta = 45°$ and 135° in all the curves. The heights indicate how hard the axis is. For example, for $E = -3.9$ MV cm^{-1} the 135° peak is lower than the 45° peak, indicating that the [$\bar{1}$10] direction is easier than [110] (the uniaxial easy axis). Consequently, the results described above confirm that our sample has a combination of two components of the magnetic anisotropy; the biaxial component and the uniaxial component. Close inspection of Fig. 2c shows that the height of the 45° peak decreases, whereas that of the 135° peak increases, when E is scanned from -3.9 to 3.9 MV cm^{-1}. The 135° peak is higher than the 45° peak at $E = 3.9$ MV cm^{-1}, showing that the uniaxial easy axis switches from [$\bar{1}$10] to [110] as E is varied. Repetition of the experiment reproduces the observed results without noticeable differences.

We now determine the magnetic anisotropy fields to establish the magnetic free energy F quantitatively and, so, obtain the dependence on E of the direction of **M** for $H = 0$. We fit the measured R_{xy} using a coherent rotation model from previous studies[16,17,20,21], in which the total magnetic free energy is as follows:

$$F = \frac{MH_B}{8}\sin^2 2\varphi + \frac{MH_U}{2}\sin^2(\varphi - 45°) - MH\cos(\varphi - \theta)$$

Here H_B and H_U are the biaxial and the uniaxial anisotropy fields, respectively, and $M = |\mathbf{M}|$. The first term corresponds to the biaxial anisotropy energy along [100] and [010], the second term to the uniaxial anisotropy energy along [110], and the third term to the Zeeman energy. In our definition, [110] is the easy axis when $H_U > 0$. By imposing the conditions $\partial F/\partial\varphi = 0$ and $\partial^2 F/\partial\varphi^2 > 0$ and using the relation $R_{xy}/R_{xy0} = \cos 2\varphi$, we fit the curves of R_{xy}/R_{xy0} versus θ to the experimental data using H_B and H_U as adjustable parameters. Our procedure reproduces the data (for example, see Fig. 2a). The values of $\mu_0 H_B$ and $\mu_0 H_U$ determined as functions of E are plotted in Fig. 3a. The value of $\mu_0 H_U$ increases with E and changes sign from negative to positive at $E \approx 2$ MV cm^{-1}, whereas $\mu_0 H_B$ remains almost independent of E, suggesting that the biaxial anisotropy constant is E-insensitive under the assumption of constant M[20] (see also Methods Summary). From the $\mu_0 H_B$ and $\mu_0 H_U$ values obtained, we then determine the dependence of φ on E at $H = 0$ (Fig. 3b). The value of φ at $E = 0$ evaluated by magnetization measurements at $H = 0$ (see Methods Summary) is shown by an open star in Fig. 3b, and is in good agreement with the value determined from the transport measurements at $H = 0$, supporting the method we used to determine φ. We see from this figure that the direction of **M** (that is, φ) can be manipulated by changing **E**. No magnetic field, spin current, mechanical strain from an external piezo actuator[7,8], or light illumination[23,24] is required.

The dependence on E of the sheet hole concentration p_{sheet} (Fig. 3c) is determined from the dependence on E of the sheet conductivity $G_{xx} = 1/R_{xx}$. Because $G_{xx} = e\mu p_{sheet} = \mu C(V_G - V_T)$, where e is the elementary charge, μ is the mobility, C is the gate capacitance and V_T is the threshold voltage, it follows that $p_{sheet} = G_{xx}/[(1/C)(dG_{xx}/dV_G)e]$. Two methods are employed to evaluate $\mu = (1/C)(dG_{xx}/dV_G)$: one uses the slope of the G_{xx}–E curve at $E < 0$ to avoid the effect of the interface states and the other uses the derivative dG_{xx}/dV_G at each point. Using the first method results in the lower limit of p_{sheet} and using the second method yields the upper limit; these are shown in Fig. 3c (for details, see Supplementary Information). Comparison of Fig. 3b and Fig. 3c shows that, in the present device, the direction of **M** can be manipulated by 10° by changing p_{sheet} from 6×10^{13} cm^{-2} (at -3.9 MV cm^{-1}) to 0.5×10^{13} cm^{-2} (averaged at 3 MV cm^{-1}), which corresponds to changing $p = p_{sheet}t$ from 1.5×10^{20} to 0.13×10^{20} cm^{-3}, where $t = 4$ nm is the channel thickness. The hole concentration corresponding to $E \approx 2$ MV cm^{-1}, where the sign change occurs in the uniaxial anisotropy, is $p = (1.6 \pm 1.3) \times 10^{19}$ cm^{-3}.

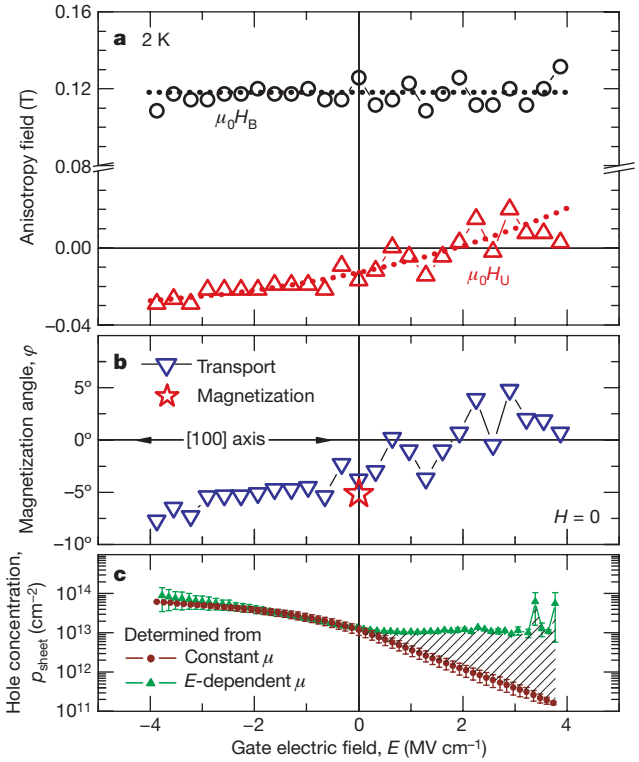


Figure 3 | Anisotropy fields, magnetization angle and sheet hole concentration as functions of electric field. **a**, Dependence on E of biaxial ($\mu_0 H_B$, open circles) and uniaxial ($\mu_0 H_U$, open triangles) anisotropy fields, at 2 K. **b**, The magnetization angle φ at $H = 0$ (triangles) as a function of E, determined from $\mu_0 H_B$ and $\mu_0 H_U$. The star was obtained from magnetization measurement, and shows good agreement with the value of φ determined from transport measurement. **c**, Dependence on E of sheet carrier concentration p_{sheet}, determined from a constant mobility (filled circles) and an E-dependent mobility (filled triangles), which give the lower and upper limits of p_{sheet}, respectively. Data are median ± s.e.m.; $n = 2$ measures (for details, see Supplementary Information).

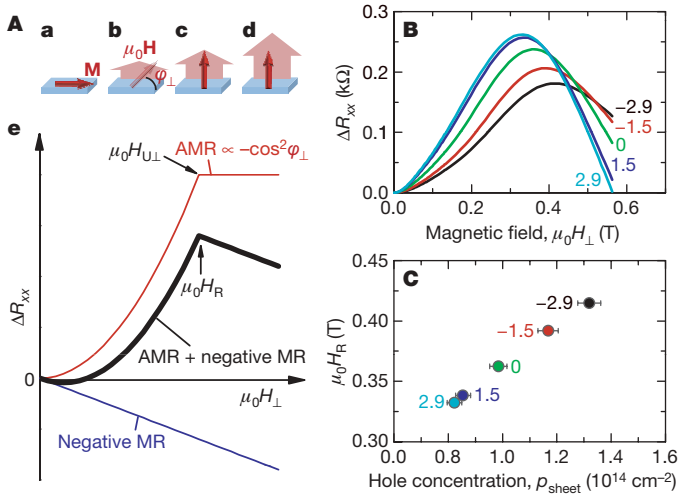


Figure 4 | Electric-field-dependent perpendicular uniaxial hard-axis anisotropy. **A**, **a**–**d**, The magnetization process; see the text. **e**, Anisotropic magnetoresistance (AMR), negative magnetoresistance (MR), and the sum of the two (thick solid line), as a function of perpendicular magnetic field, $\mu_0 H_\perp$. The magnetic field at which ΔR_{xx} (the difference between $R_{xx}(H)$ and $R_{xx}(H = 0)$) peaks, $\mu_0 H_R$, is a good measure of the perpendicular uniaxial anisotropy field, $\mu_0 H_{U\perp}$. **B**, Dependence on $\mu_0 H_\perp$ of ΔR_{xx}, for five different values of E (labelling the plots in units of MV cm^{-1}), at 5 K. $\mu_0 H_R$ becomes lower when a positive E is applied and becomes higher when negative E is applied. **C**, $\mu_0 H_R$ as a function of the sheet hole concentration, p_{sheet}, which data (median ± s.e.m., $n = 2$; labelled with E as in **B**) are determined from the average of both sweep-up and sweep-down curves of the E dependence of the sheet conductivity.

Finally, we show the E-dependent perpendicular hard-axis anisotropy field $\mu_0 H_{U\perp}$, which makes it possible to control the out-of-plane rotation of **M** using **E** under a constant, non-zero out-of-plane μ_0**H**. We evaluate the dependence of $\mu_0 H_{U\perp}$ on E from the anisotropic magnetoresistance (AMR)[25,26] using the dependence of R_{xx} on $\mu_0 H_\perp$, the plane-perpendicular component of μ_0**H**. Figure 4A, e shows a sketch of the dependence of AMR on $\mu_0 H_\perp$, the contribution of negative magnetoresistance[26,27], and the sum of the two. The vertical axis, ΔR_{xx}, is the difference between $R_{xx}(H_\perp)$ and $R_{xx}(H = 0)$. As depicted in Fig. 4A, a, for $H = 0$ **M** lies in the plane. As $\mu_0 H_\perp$ increases, the perpendicular component of **M** gradually increases (Fig. 4A, b). R_{xx} also increases, because of the dominant contribution of AMR (proportional to $-\cos^2\varphi_\perp$, where $\varphi_\perp$ is the angle between **M** and the plane). Further increase of $\mu_0 H_\perp$ to $\mu_0 H_{U\perp}$ aligns **M** parallel to $\mu_0 H_\perp$, where AMR saturates (Fig. 4A, c) and beyond which R_{xx} is reduced by the negative magnetoresistance (Fig. 4A, d). The magnetic field at the maximum of R_{xx}, $\mu_0 H_R$, is thus a good measure of $\mu_0 H_{U\perp}$. Figure 4B shows the dependence of ΔR_{xx} on $\mu_0 H_\perp$ for five different values of E at 5 K, for device B. We see the shift of $\mu_0 H_R$ to lower or higher values under positive or negative values of E, respectively, indicating the change in $\mu_0 H_{U\perp}$. In Fig. 4C, we plot $\mu_0 H_R$ as a function of p_{sheet}. The results demonstrate that **E** can be used to tilt the direction of **M** under a finite $\mu_0 H_\perp$ less than $\mu_0 H_{U\perp}$.

The observed manipulation of **M** by **E** is based on the control of magnetic anisotropy through changing p. The p dependence of the magnetic anisotropies has been predicted using the p–d Zener model Hamiltonian approach[28–30]. There, the anisotropic valence band structure, which depends on the magnitude and the direction of **M** as well as the value of p, is responsible for the anisotropy; the model correctly reproduces the strain dependence of the magnetic anisotropy[30]. According to the model, the biaxial magnetic anisotropy is p dependent, whereas the present study shows that the uniaxial magnetic anisotropy is more sensitive to the change of p. Additional physics might have to be included in the existing model to explain our results.

Work for the future includes the switching of **M** between stable directions, which adds a non-volatile device function to the present device, where the application of **E** is required only during the control of the direction of **M**. Our calculation backed up by micromagnetic simulation (Supplementary Information) shows that when a set of conditions (one of which is the sign change of $\mu_0 H_U$ as **E** is varied, observed here) is met the switching takes place. This opens a route to a totally new scheme of non-volatile device operation, without resorting to the use of a magnetic field, spin current or mechanical stress.

METHODS SUMMARY

Sample growth. A 4-nm-thick $Ga_{0.9}Mn_{0.1}As$ layer with a 2-nm GaAs cap layer was grown at 200 °C by molecular beam epitaxy. The buffer layer underneath consists of 4-nm GaAs/30-nm $Al_{0.8}Ga_{0.2}As$/30-nm GaAs, which was grown on a semi-insulating GaAs (001) substrate.

Magnetization measurements. The dependence of the remnant magnetization, M_r, on temperature, T, was measured after the application of an external magnetic field $\mu_0 H = 0.1$ T at 2 K, using a superconducting quantum interference device magnetometer and a sample with large area (~16 mm^2) cleaved from the same wafer as devices A and B. The magnetization angle φ, measured from [100], can be determined from $\varphi = 45° - \tan^{-1}(M_r^{[\bar{1}10]}/M_r^{[110]})$, where $M_r^{[\bar{1}10]}$ and $M_r^{[110]}$ are the respective values of M_r along [$\bar{1}$10] and [110]. From measurement of M_r as a function of T, the Curie temperature of this sample was determined to be 46 K, which shows good agreement with the value determined from the maximum of the R_{xx}–T curve at $E = 0$. The assumption that M is constant and independent of E is based on the fact that the working temperature of 2 K in the present experiment is much lower than the Curie temperature of 46 K, even when it is modulated by ±10 K by E.

Received 26 May; accepted 22 July 2008.

1. Slonczewski, J. C. Current-driven excitation of magnetic multilayers. *J. Magn. Magn. Mater.* **159,** 1–7 (1996).
2. Berger, L. Emission of spin waves by a magnetic multilayer traversed by a current. *Phys. Rev. B* **54,** 9353–9358 (1996).
3. Hosomi, M. *et al.* A novel nonvolatile memory with spin torque transfer magnetization switching spin-RAM. *IEDM Tech. Dig.* **473,** 459–462 (2005).
4. Kawahara, T. *et al.* in *ISSCC 2007* (eds Amiri, M *et al.*) 480–481 (IEEE Internat. Solid-State Circuits Conf., IEEE, 2007).
5. Schröder, K. Design parameters of a 3-dimensional ultrasonic pulse controlled memory device with single domain coherently magnetized cobalt, iron and nickel particles in a non-magnetic matrix. *IEEE Trans. Magn.* **10,** 567–570 (1974).
6. Novosad, V. *et al.* Novel magnetostrictive memory device. *J. Appl. Phys.* **87,** 6400–6402 (2000).
7. Goennenwein, S. T. B. *et al.* Piezo-voltage control of magnetization orientation in a ferromagnetic semiconductor. *Phys. Status Solidi (RRL)* **2,** 96–98 (2008).
8. Overby, M., Chernyshov, A., Rokhinson, L. P., Liu, X. & Furdyna, J. K. GaMnAs-based hybrid multiferroic memory device. *Appl. Phys. Lett.* **92,** 192501 (2008).
9. Tokura, Y. Multiferroics as quantum electromagnets. *Science* **312,** 1481–1482 (2006).
10. Eerenstein, W., Mathur, N. D. & Scott, J. F. Multiferroic and magnetoelectric materials. *Nature* **442,** 759–765 (2006).
11. Chu, Y.-H. *et al.* Electric-field control of local ferromagnetism using a magnetoelectric multiferroic. *Nature Mater.* **7,** 478–482 (2008).
12. Ohno, H. *et al.* Electric-field control of ferromagnetism. *Nature* **408,** 944–946 (2000).
13. Chiba, D., Yamanouchi, M., Matsukura, F. & Ohno, H. Electrical manipulation of magnetization reversal in a ferromagnetic semiconductor. *Science* **301,** 943–945 (2003).
14. Chiba, D., Matsukura, F. & Ohno, H. Electric-field control of ferromagnetism in (Ga,Mn)As. *Appl. Phys. Lett.* **89,** 162505 (2006).
15. Shen, A. *et al.* Epitaxy of (Ga,Mn)As, a new diluted magnetic semiconductor based on GaAs. *J. Cryst. Growth* **175–176,** 1069–1074 (1997).
16. Tang, H. X., Kawakami, R. K., Awschalom, D. D. & Roukes, M. L. Giant planar Hall effect in epitaxial (Ga,Mn)As devices. *Phys. Rev. Lett.* **90,** 107201 (2003).
17. Welp, U., Vlasko-Vlasov, V. K., Liu, X., Furdyna, J. K. & Wojtowicz, T. Magnetic domain structure and magnetic anisotropy in $Ga_{1-x}Mn_xAs$. *Phys. Rev. Lett.* **90,** 167206 (2003).
18. Liu, X., Sasaki, Y. & Furdyna, J. K. Ferromagnetic resonance in $Ga_{1-x}Mn_xAs$: Effects of magnetic anisotropy. *Phys. Rev. B* **67,** 205204 (2003).
19. Sawicki, M. *et al.* In-plane uniaxial anisotropy rotations in (Ga,Mn)As thin films. *Phys. Rev. B* **71,** 121302 (2005).
20. Wang, K.-Y. *et al.* Spin reorientation transition in single-domain (Ga,Mn)As. *Phys. Rev. Lett.* **95,** 217204 (2005).
21. Hamaya, K. *et al.* Magnetic anisotropy switching in (Ga,Mn)As with increasing hole concentration. *Phys. Rev. B* **74,** 045201 (2006).
22. Pappert, K. *et al.* Transport characterization of the magnetic anisotropy of (Ga,Mn)As. *Appl. Phys. Lett.* **90,** 062109 (2007).
23. Oiwa, A., Mitsumori, Y., Moriya, R., Słupinski, T. & Munekata, H. Effect of optical spin injection on ferromagnetically coupled Mn spins in the III–V magnetic alloy semiconductor (Ga,Mn)As. *Phys. Rev. Lett.* **88,** 137202 (2002).
24. Hashimoto, Y., Kobayashi, S. & Munekata, H. Photoinduced precession of magnetization in ferromagnetic (Ga,Mn)As. *Phys. Rev. Lett.* **100,** 067202 (2008).
25. Hayashi, T. *et al.* Anisotropy and Barkhausen jumps in diluted magnetic semiconductor (Ga,Mn)As. *Physica B* **284–288,** 1175–1176 (2000).
26. Matsukura, F., Sawicki, M., Dietl, T., Chiba, D. & Ohno, H. Magnetotransport properties of metallic (Ga,Mn)As films with compressive and tensile strain. *Physica E* **21,** 1032–1036 (2004).
27. Matsukura, F., Ohno, H., Shen, A. & Sugawara, Y. Transport properties and origin of ferromagnetism in (Ga,Mn)As. *Phys. Rev. B* **57,** R2037–R2040 (1998).
28. Dietl, T., Ohno, H., Matsukura, F., Cibert, J. & Ferrand, D. Zener model description of ferromagnetism in zinc-blend magnetic semiconductors. *Science* **287,** 1019–1022 (2000).
29. Abolfath, M., Jungwirth, T., Brum, J. & MacDonald, A. H. Theory of magnetic anisotropy in $III_{1-x}Mn_xV$ ferromagnets. *Phys. Rev. B* **63,** 054418 (2001).
30. Dietl, T., Ohno, H. & Matsukura, F. Hole-mediated ferromagnetism in tetrahedrally coordinated semiconductors. *Phys. Rev. B* **63,** 195205 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We acknowledge discussion with M. Shirai. This work was supported in part by Grant-in-Aids from MEXT/JSPS, the GCOE Program at Tohoku University, the Research and Development for Next-Generation Information Technology Program (MEXT), and a Research Fellowship from JSPS.

Author Contributions D.C. and H.O. planned and supervised the study. D.C. grew the samples, made the devices, collected and analyzed data, and calculated magnetization switching. M.S. and F.M. investigated the magnetization characteristics of (Ga,Mn)As layers. Y. Nishitani investigated the transport characteristics of devices for the transport measurements and determined the dielectric constant of the ZrO_2 gate insulators. Y. Nakatani performed the micromagnetic simulation. D.C., F.M. and H.O. wrote the manuscript. All authors discussed the results.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to H.O. (ohno@riec.tohoku.ac.jp).

LETTERS

Observed and modelled stability of overflow across the Greenland–Scotland ridge

Steffen M. Olsen[1], Bogi Hansen[2], Detlef Quadfasel[3] & Svein Østerhus[4]

Across the Greenland–Scotland ridge there is a continuous flow of cold dense water, termed 'overflow', from the Nordic seas to the Atlantic Ocean[1]. This is a main contributor to the production of North Atlantic Deep Water[2] that feeds the lower limb of the Atlantic meridional overturning circulation, which has been predicted to weaken as a consequence of climate change[3,4]. The two main overflow branches pass the Denmark Strait and the Faroe Bank channel. Here we combine results from direct current measurements in the Faroe Bank channel[5] for 1995–2005 with an ensemble hindcast experiment[6] for 1948–2005 using an ocean general circulation model. For the overlapping period we find a convincing agreement between model simulations and observations on monthly to interannual timescales. Both observations and model data show no significant trend in volume transport. In addition, for the whole 1948–2005 period, the model indicates no persistent trend in the Faroe Bank channel overflow or in the total overflow transport, in agreement with the few available historical observations. Deepening isopycnals in the Norwegian Sea have tended to decrease the pressure difference across the Greenland–Scotland ridge[7], but this has been compensated for by the effect of changes in sea level. In contrast with earlier studies[7,8], we therefore conclude that the Faroe Bank channel overflow, and also the total overflow, did not decrease consistently from 1950 to 2005, although the model does show a weakening total Atlantic meridional overturning circulation as a result of changes south of the Greenland–Scotland ridge.

According to observations, the total volume transport of overflow water across the Greenland–Scotland ridge (GSR)[1] is about 6 Sv (1 Sv (sverdrup) = 10^6 m^3 s^{-1}) (Fig. 1a). Half of this passes through the Denmark Strait; the other half crosses the ridge east of Iceland through several passages, mainly through the Faroe Bank channel (FBC). At the sill of the FBC, the overflow current is only about 10 km wide (Fig. 1b), which is comparable to the baroclinic Rossby radius. This permits accurate monitoring of volume transport with limited instrumentation[5]. Since 1995 an acoustic Doppler current profiler has been continuously moored in the centre of the channel over the sill (site B; Fig. 1b). Combined with an extensive suite of short-term experiments and hydrographic monitoring, this has permitted the generation of a time series of the overflow volume transport of high accuracy[5], with a mean transport of 2.1 Sv (Fig. 2a). In addition to the FBC overflow, the eastern overflow includes a total of 1.3 Sv of overflow across the Wyville Thomson ridge (0.3 Sv (ref. 9)) and the Iceland–Faroe ridge (1 Sv (refs 10–12)). These shallow overflows are intermittent and are dominated by synoptic-scale processes rather than the more stable forcing maintaining the overflows in the Denmark Strait and the FBC. No long-term direct transport observations exist; we therefore have no information on flow variability on timescales above seasonal.

A second independent estimate of the overflow transport stems from an ensemble hindcast simulation[6] for the period 1948–2005 with the use of a global coupled ocean–sea-ice model[13] forced by atmospheric reanalysis data and observed Arctic river discharges[6]

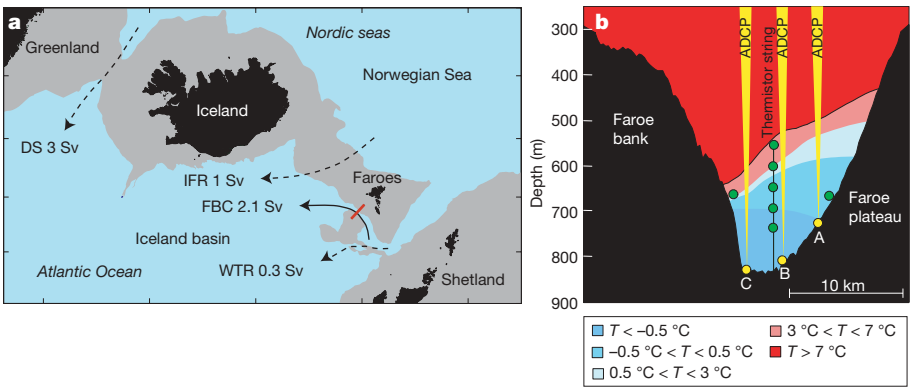


Figure 1 | The GSR overflow branches. **a**, Map of the GSR, with overflow branches indicated by arrows. Numbers give the typical volume transports of each branch[5,9–12,20]. The solid arrow shows the FBC overflow; dashed arrows indicate the Wyville Thomson ridge (WTR), the Iceland–Faroe ridge (IFR) and Denmark Strait (DS) overflows. The grey area is shallower than 750 m. The red line indicates the monitoring section in the FBC at the sill. **b**, On this section, instruments have been moored for periods of up to 10 years. Yellow circles indicate the positions of moored acoustic Doppler current profilers with sound beams illustrated by yellow cones. Green circles show the locations of moored temperature sensors. Background colours show the typical temperature distribution on the section.

[1]Danish Meteorological Institute, Lyngbyvej 100, 2100 Copenhagen, Denmark. [2]Faroese Fisheries Laboratory, Box 3051, FO-110 Tórshavn, Faroe Islands. [3]Universität Hamburg, Zentrum für Meeres- und Klimaforschung, Bundesstrasse 53, D-20146 Hamburg, Germany. [4]Bjerknes Centre for Climate Research, University of Bergen, NO-5014 Bergen, Norway.

(see Methods). Individual exchange branches of Atlantic, overflow and polar waters are diagnosed from the simulation by using temperature and salinity characteristics. The ensemble mean climatology of exchanges[6], including a total overflow across the GSR of 6.0 Sv and a net Atlantic inflow of 8.7 Sv, compares favourably with existing observations[1,14]. The model does not separate the various overflows east of Iceland, owing to limited grid resolution. It provides a time series of total eastern overflow with an ensemble mean transport of 3.4 Sv for the period 1995–2005. This is consistent with observations of the combined eastern overflow transport (Fig. 1a). Subtracting the mean overflow across the shallow ridges (1.3 Sv) from the total, we generated a time series of computed FBC overflow (Fig. 2a).

The correspondence between this time series of FBC overflow and the observational series for 1995–2005 is remarkable and can be verified independently to arise from identical seasonality and highly correlated monthly and interannual variability (see Methods). On monthly scales, the correlation coefficient calculated by using deseasoned data is 0.72 ($P < 10^{-7}$), whereas annual mean data (Fig. 2a) show an even higher correlation of 0.92 ($P < 10^{-4}$). The statistical analysis confirms that this is extremely unlikely to occur by chance. No individual ensemble member fits the observations better than the ensemble mean (Supplementary Fig. 1), in which internal model variability is practically eliminated (Supplementary Fig. 2). Internal variability may have different sources but is not examined in detail (see Methods). Clearly, however, ensemble averaging allows a very accurate representation of the FBC overflow, which is a strongly forced system with relatively small internal variability.

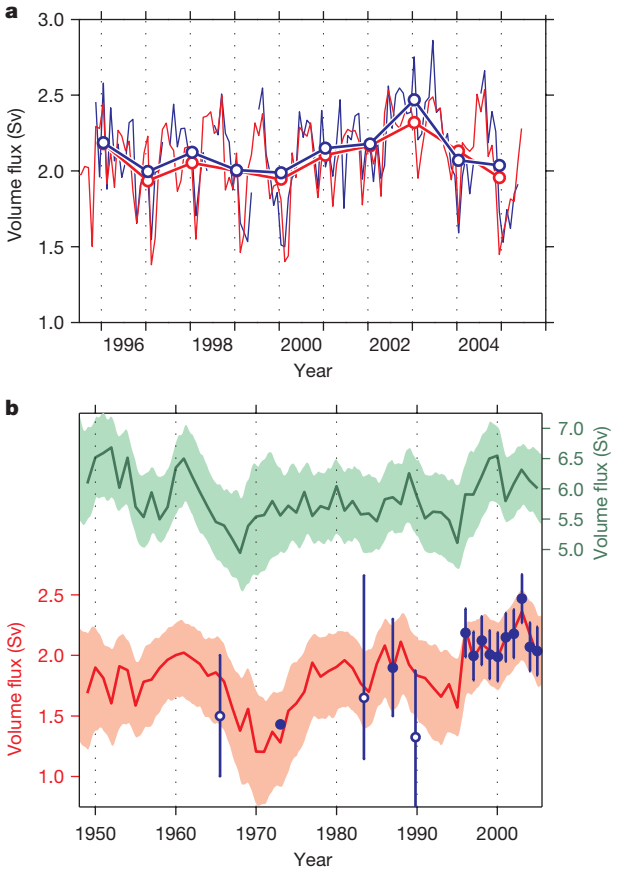


Figure 2 | Observed and computed deep overflow. a, Monthly mean observations of the FBC overflow (thin blue lines) and a time series (thin red lines) computed by subtracting 1.3 Sv from the ensemble mean modelled eastern overflow in the instrumental period. Annual mean values (connected circles) are included. Here, data are averaged for months where observations exist (July to May). **b**, Annual FBC overflow transport (red) and total overflow transport across the GSR (green) for 1948–2005, computed on the basis of ensemble means. The ensemble spread between members is represented by coloured shading (2σ). Blue filled circles show all published, long-term (weeks to months) observations of the FBC overflow transport[5,16,24,25], and open circles show direct, near-synoptic transport estimates[26–28]. Error bars show published error estimates for the respective observations when available. The observed FBC overflow series since 1995 are defined solely from the velocity field[5]. For the model series, additional temperature and salinity constraints have been applied[6]. These definitions are internally consistent and comparable to procedures used to obtain overflow transport estimates before 1995.

The trends in both series are equal, weak and not significantly different from zero (0.01 ± 0.02 Sv per decade). The highly significant correlations between two independently produced time series verify the accuracy of both series conclusively. This also implies that the FBC overflow has not changed consistently during the 1995–2005 period. Encouraged by the excellent agreement between model and observations, we used the model data back to 1948 to estimate long-term variations in both the FBC and the total overflow transports (Fig. 2b). Here, the estimated trend of the FBC overflow is found to be positive but again barely distinguishable from zero (0.05 ± 0.02 Sv per decade). For the early period, direct estimates of FBC overflow are few and not of the same proven quality as after 1995, but they do agree with the model time series (Fig. 2b).

This result of no detectable long-term trend in the overflow contradicts earlier reported weakening[7] of the FBC overflow and is hard to reconcile with measurements indicating a marked decrease[8] in the Atlantic meridional overturning circulation (AMOC) at 25° N, results that have also been questioned by other studies[15]. To understand this discrepancy we performed an analysis of the forcing of the overflow. Theoretical arguments[5,7,16–20] imply that the FBC overflow transport is related to the pressure difference (ΔP) between both sides of the GSR at the depth of the overflow core. This includes a barotropic component (ΔP_{trop}) determined by the sea level and a baroclinic component (ΔP_{clin}) given by the density field ($\Delta P = \Delta P_{\mathrm{trop}} + \Delta P_{\mathrm{clin}}$). The model shows an almost linear relationship between the overflow transport and the total pressure difference (ΔP) (Fig. 3b). Local and remote processes such as convection, mixing and the circulation in the basins north and south of the ridge contribute to determining the pressure difference; however, once given, it specifies the FBC overflow.

According to the model, the baroclinic component has an overall decreasing trend between 1948 and 2005 (Fig. 3a), consistent with earlier reports[7]. In contrast, the barotropic pressure gradient shows an increasing trend of equal magnitude, so the pressure difference at depth does not have a significant trend. This compensating effect of barotropic and baroclinic contributions to the pressure difference is a striking feature of the time series on timescales from a few years to decades (Fig. 3a). Wind stress affects the strength of the open ocean gyres and, through geostrophy, determines the local sea-level topography including the barotropic pressure difference across the GSR[21]. With some time lag, the density field adjusts to compensate partly for barotropic variations (Fig. 3a). A closer analysis of the pressure difference variations in the model (Fig. 4) reveals that most of the variation originates in the Nordic seas. The changes are of basin-wide scale, and the strong tendency of compensation (Fig. 3a) is thus independent of the specific positions chosen to calculate the cross-ridge pressure variations. Since 1948 the model sea level over most of the Nordic seas showed a positive trend (Fig. 4a), although not monotonic (Fig. 3a); however, adjustment of the density field (Fig. 4b) by lowering isopycnal surfaces led to the maintenance of a relatively stable overflow on timescales above a few years. This timescale of adjustment is quantified in a conceptual framework and explains the main features of Fig. 3a (see Methods).

The conclusion of a relatively stable FBC overflow since 1948 is based on the model results and is corroborated by early, direct estimates of the transport (Fig. 2b). In addition, variations in the pressure difference that forces the overflow can be further validated against observations. Open-ocean sea-level gradients derived from

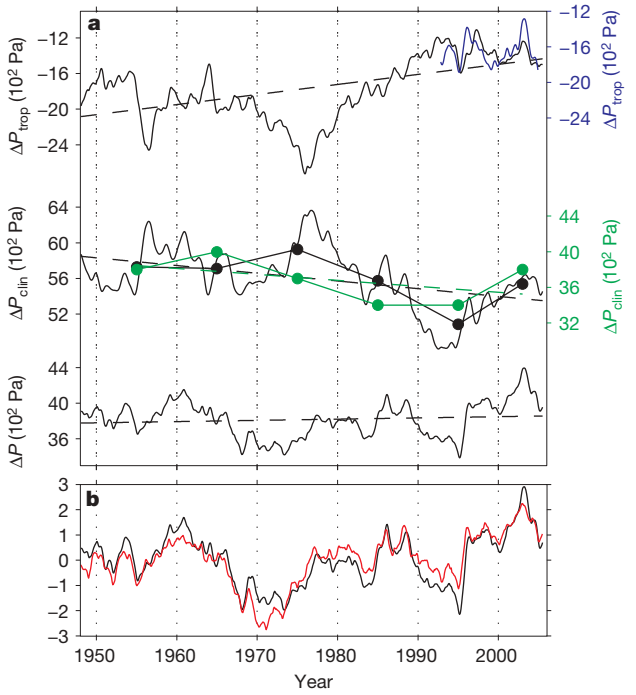


Figure 3 | Pressure forcing of the FBC overflow. **a**, Modelled (black) barotropic (ΔP_{trop}), baroclinic (ΔP_{clin}) and total (ΔP) pressure differences between both ends of the channel at the depth of the overflow core (750 m), evaluated from two remote positions in the Norwegian Sea and the Iceland basin (north minus south; for positions see Fig. 4). For ΔP_{clin}, decadal averages are included (circles) and compared with observed ΔP_{clin} (green) on the basis of hydrographic profiles extracted from the quality-controlled International Council for the Exploration of the Seas (ICES) database[29] near the two end points; note the different scales. For ΔP_{trop}, the modelled variability is compared with the observed gradient derived from altimetry (blue) for the period 1992–2005. Dashed lines represent linear regression fits to the model time series (black) and to the observed ΔP_{clin} (green). **b**, The direct pressure forcing of the modelled overflow transport is illustrated by comparing the evolution of the normalized pressure difference (black) and computed overflow transport (red). The correlation coefficient between pressure and flow is 0.90. The annual cycle is removed from the monthly model and altimetry data presented, and the time series have subsequently been low-pass filtered with a forward–reverse filter with a cutoff period of 12 months.

satellite altimetry since 1992 match the temporal evolution of the barotropic forcing simulated by the model fairly well (Fig. 3a) but the period is too short for a trend analysis. Hydrographic observations span the whole period, but they are sparse and inhomogeneous in time and space and do not permit the resolution of interannual variability. However, comparison between the model-derived and observational baroclinic pressure gradients does yield a similar level of inter-decadal variability and nearly identical, negative trends (Fig. 3a). This confirms that the model does not underestimate the decline in the baroclinic forcing of the overflow during the past five decades[7]. The comparison reveals an offset of about 20×10^2 Pa between model and observations (Fig. 3a). This discrepancy is intrinsic to the limited grid resolution of the model, which apparently overestimates the baroclinic (and hence total) cross-ridge pressure difference needed to drive a realistic eastern overflow. As long as this offset is not too large, it will not affect the shorter-term and longer-term variability of the transports, as indeed is suggested by the good correspondence between observations and the model (Fig. 2).

The compensation between barotropic and baroclinic pressure gradients has consequences for the future monitoring of the overflows. Monitoring of the hydrographic structure and sea-surface height alone[22] is faced with the problem of disentangling the relatively small residual signal of interest as a difference of two larger, partly compensating, signals. At present, there is therefore no alternative to monitoring by direct current measurements.

In this paper we argue that the high correlation between observed and computed FBC overflow transports provides strong evidence that the model output is realistic. Strictly, this has only been proven for the FBC overflow in the period 1995–2005, but model validity for the total overflow and for the whole period 1948–2005 are supported by the available observations: overflow measurements in the Denmark Strait[20], historic FBC overflow estimates (Fig. 2b), satellite altimetry (Fig. 3a) and hydrographic observations (Fig. 3a).

The combined evidence from observations and model results indicates that neither the FBC overflow nor the total overflow transports had any consistent trends from 1948 to 2005. This is in conflict with the claim[8] that the overflow contribution to the AMOC at 25° N weakened by 50% during this period. In the present climate, the overflows supply about one-third to the volume transport of the deep branch of the AMOC[23]. Therefore any variations in the AMOC cannot be related to overflow variability but instead have to be produced by processes south of the GSR such as entrainment and Labrador Sea convection, each of which contribute about one-third to the AMOC.

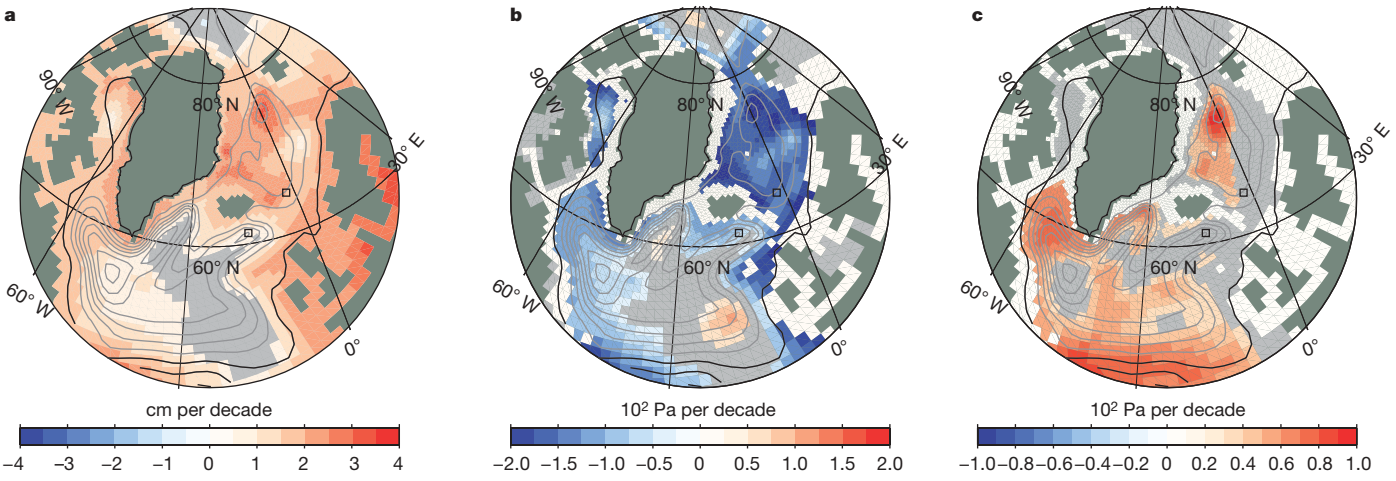


Figure 4 | Modelled long-term changes in the North Atlantic. Decadal tendencies calculated for the full period 1948–2005 are shown for sea-level (**a**), baroclinic (**b**) and total (**c**) pressure at a level representing the core of the FBC deep overflow (750 m). Regions where the trend is not significant are shown in grey. The time-mean barotropic stream-function is depicted with cyclonic circulation indicated by grey lines. Transport between contours is 4 Sv. The squares mark the grid points between which pressure differences (Fig. 3) were calculated.

METHODS SUMMARY

We exploit an ensemble global ocean hindcast simulation[6] with 26 members, using a coupled ocean–sea-ice model, forced by reanalysis data. The model configuration has coarse grid spacing, although enhanced horizontal resolution is achieved in the Nordic seas. The model grid does not resolve the various shallow overflows east of Iceland but has one channel that carries the total eastern overflow. The experiment is designed[6] to isolate the forced ocean climate history by construction of the ensemble mean, thereby reducing the uncertainties arising from the initial state and internal model variability. The ensemble spread decreases over time[6] (Fig. 2b), and short-term variability is largely unaffected by averaging (Supplementary Fig. 2). The resulting mean time series fits the observations extremely well (Fig. 2a), better than any individual member (Supplementary Fig. 1). We test the significance of our correlations between computed and observed FBC overflow and find the probability of obtaining a correlation coefficient of 0.72 for monthly, de-seasoned variability by chance to be less than 10^{-7} and the annual mean correlation of 0.92 to be less than 10^{-4}.

Compensation between barotropic and baroclinic changes (Fig. 3) is explained conceptually by using the linear relation between variations in simulated overflow transport and pressure difference between the relatively quiescent basin downstream of the channel and the Norwegian basin, upstream of the channel. An increase in sea level here increases the difference in barotropic pressure, and the associated increase in overflow transport implies an increased removal of overflow water from the basin. This causes the isopycnal surfaces to deepen, reducing the baroclinic pressure difference. If the Norwegian basin is approximated with a two-layer system, the timescale associated with this compensating effect is about three years.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 18 January; accepted 31 July 2008.

1. Saunders, P. M. in *Ocean Circulation and Climate: Observing and Modelling the Global Ocean* (eds Siedler, G., Church, J. & Gould, J.) 401–417 (Academic, 2001).
2. Dickson, R. R. & Brown, J. The production of North Atlantic Deep Water: Sources, rates, and pathways. *J. Geophys. Res.* **99,** 12319–12342 (1994).
3. Gregory, J. M. *et al.* A model intercomparison of changes in the Atlantic thermohaline circulation in response to increasing atmospheric CO_2 concentrations. *Geophys. Res. Lett.* **32,** L12703, doi:10.1029/2005GL023209 (2005).
4. Meehl, G. A. *et al.* in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Solomon, S. *et al.*) 747–846 (Cambridge Univ. Press, 2007).
5. Hansen, B. & Østerhus, S. Faroe Bank Channel overflow 1995–2005. *Prog. Oceanogr.* **75,** 817–856 (2007).
6. Olsen, S. M. & Schmith, T. North Atlantic–Arctic Mediterranean exchanges in an ensemble hindcast experiment. *J. Geophys. Res.* **112,** C04010, doi:10.1029/2006JC003838 (2007).
7. Hansen, B., Turrel, W. R. & Østerhus, S. Decreasing overflow from the Nordic seas into the Atlantic Ocean through the Faroe Bank channel since 1950. *Nature* **441,** 927–930 (2001).
8. Bryden, H. L., Longworth, H. R. & Cunningham, S. A. Slowing of the Atlantic meridional overturning circulation at 25° N. *Nature* **438,** 655–657 (2005).
9. Sherwin, T. J., Griffiths, C. R., Inall, M. E. & Turrell, W. R. Quantifying the overflow across the Wyville Thomson Ridge into the Rockall Trough. *Deep-Sea Res.* **55,** 396–404 (2008).
10. Hermann, F. The T-S diagram analysis of the water masses over the Iceland–Faroe Ridge and in the Faroe Bank Channel (Overflow '60). *Rapp. P.-V. Réun. Cons. Int. Explor. Mer* **157,** 139–149 (1967).
11. Meincke, J. On the distribution of low salinity intermediate waters around the Faroese. *Dt. Hydrogr. Z.* **31,** 50–64 (1978).
12. Perkins, H., Hopkins, T. S., Malmberg, S.-A., Poulain, P.-M. & Warn-Varnas, A. Oceanographic conditions east of Iceland. *J. Geophys. Res.* **103,** 21531–21542 (1998).
13. Marsland, S. J., Haak, H., Jungclaus, J. H., Latif, M. & Röske, F. The Max-Plank Institute global ocean/sea ice model with orthogonal curvilinear coordinates. *Ocean Model.* **5,** 91–127 (2003).
14. Østerhus, S., Turrell, W. R., Jónsson, S. & Hansen, B. Measured volume, heat, and salt fluxes from the Atlantic to the Arctic Mediterranean. *Geophys. Res. Lett.* **32,** L07603, doi:10.1029/2004GL022188 (2005).
15. Cunningham, S. A. *et al.* Temporal variability of the Atlantic meridional overturning circulation at 26.5°N. *Science* **317,** 935–937 (2007).
16. Crease, J. The flow of Norwegian Sea Water through the Faroe Bank Channel. *Deep-Sea Res.* **12,** 143–150 (1965).
17. Whitehead, J. A. Topographic control of oceanic flows in deep passages and straits. *Rev. Geophys.* **36,** 423–440 (1998).
18. Girton, J. B., Pratt, L. J., Sutherland, D. A. & Price, J. F. Is the Faroe Bank Channel overflow hydraulically controlled? *J. Phys. Oceanogr.* **36,** 2340–2349 (2006).
19. Käse, R. H. A Riccati model for Denmark Strait overflow variability. *Geophys. Res. Lett.* **33,** L21S09, doi:10.1029/2006GL026915 (2006).
20. Dickson, B. *et al.* in *Arctic–Subarctic Ocean Fluxes: Defining the Role of the Northern Seas in Climate* (eds Dickson, R. R., Meincke, J. & Rhines, P.) 443–474 (Springer, 2008).
21. Jakobsen, P. K., Ribersgaard, M. H., Quadfasel, D., Schmith, T. & Hughes, C. W. The near-surface circulation in the northern North Atlantic as inferred from drifter data: variability from the meso-scale to interannual. *J. Geophys. Res.* **108,** 3251, doi:10.1029/2002JC001554 (2003).
22. Köhl, A., Käse, R. & Stammer, D. Causes of changes in the Denmark Strait Overflow. *J. Phys. Oceanogr.* **37,** 1678–1696 (2007).
23. Hansen, B., Østerhus, S., Quadfasel, D. & Turrell, W. Already the day after tomorrow? *Science* **305,** 953–954 (2004).
24. Saunders, P. M. Cold outflow from the Faroe Bank Channel. *J. Phys. Oceanogr.* **20,** 29–43 (1990).
25. Dooley, H. D. & Meincke, J. Circulation and water masses in the Faroese Channels during Overflow '73. *Dt. Hydrogr. Z.* **34,** 41–54 (1981).
26. Sætre, R. *Report on the Norwegian investigations in the Faeroe Channel 1964–65* (NATO Subcommittee on Oceanographic Research, Tech. Rep. No. 38, 1967).
27. Borenäs, K. A. & Lundberg, P. A. On the deep-water flow through the Faroe Bank Channel. *J. Geophys. Res.* **93,** 1281–1292 (1988).
28. Saunders, P. M. Combining hydrographic and shipborne ADCP measurements. *Deep-Sea Res.* **39,** 1412–1427 (1992).
29. International Council for the Exploration of the Seas. ICES Oceanogr. Database Services ⟨http://www.ices.dk/ocean/⟩ (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was partly supported by the European Commission as part of the MOEN (Meridional Overturning Exchange with the Nordic Seas) project, the Danish Environmental Ministry through the Danish Cooperation for Environment in the Arctic (DANCEA) programme, the Nordic Council in the project Arctic–Atlantic Exchanges (ARATEX), the German Federal Ministry of Education and Research (BMBF) in the North Atlantic Project, and the Research Council of Norway in the projects BIAC (Bipolar Atlantic Thermohaline Circulation) and POCAHONTAS (Polar Climate and Heat Transport). The altimeter products used were produced by Ssalto/Duacs and distributed by Aviso with support from CNES (http://aviso.oceanobs.com).

Author Contributions All authors contributed equally to this work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to S.M.O. (smo@dmi.dk).

METHODS

Model and ensemble approach. We exploit an ensemble global ocean hindcast simulation[6] with 26 members, using a coupled ocean–sea-ice model, the Max Planck Institute Ocean Model[13], forced by NCEP reanalysis data[30]. The particular model configuration has on average a coarse grid spacing, although enhanced horizontal resolution is achieved in the Nordic seas by displacing the model North Pole. The model grid does not resolve the various shallow overflows east of Iceland; it has one channel that carries the total eastern overflow. No restoring of ocean properties is used, but to account for missing run-off and to correct for model deficits a static correction of the evaporation–precipitation field is applied[6]. The numerical experiment has been described in detail elsewhere[6].

From an ensemble simulation, there are two basic approaches for achieving a good estimate of state variables and derived quantities. Ensemble members with optimal initial conditions may well represent real ocean characteristics and one might seek to identify these members. This is not necessarily the case for a forced experiment using a decoupled model system and was not the strategy pursued here for the ocean hindcast. Furthermore, model noise is introduced by the discrete time-stepping in combination with several highly nonlinear parameterizations of unresolved physical processes in the discretized numerical model of the mathematical governing (primitive) equations, and no single simulation can be expected to reflect directly observed changes. Hence, by constructing the ensemble mean, the experimental design[6] allows us to isolate the forced ocean variability and strongly reduces the uncertainty and noise arising from the initial state and internal model variability[31]. Although noise is expected, it is not certain to what extent the noise interacts with the forced variability, and ensemble mean data should be used with caution. A detailed analysis of the characteristics of the internal variability has not been pursued because we find that the ensemble spread of key variables shows a clear decrease in time[6] (Fig. 2b) and that the level of short-term variability is largely unaffected (Supplementary Fig. 2). In addition, convergence of the ensemble mean statistics for increasing ensemble size initially follows the scaling of an infinite synthetic ensemble with random noise added to the ensemble mean (Supplementary Fig. 2), which indicates negligible interaction between internal and forced variability. In corroboration, the ensemble mean time series fits the observations extremely well (Fig. 2a), better than any individual member (Supplementary Fig. 1). The correspondence is not really surprising, confirming that the FBC overflow is strongly forced, in the model and in nature. Model internal variability present in the individual members' estimates is strongly reduced in the ensemble mean (Supplementary Fig. 2). Internal model variability is 0.21 Sv, estimated as the mean standard deviation between ensemble members and the mean (σ_A; Supplementary Fig. 2), whereas the model-generated variability remaining after averaging is only about 0.04 Sv.

Statistical significance and implications. Standard formulae are available for testing the significance of our correlations between computed and observed FBC overflow. When comparing the two time series, one has to account for seasonality and autocorrelation, which reduces the degrees of freedom. For the overlapping period 1995–2005, we identify near-identical seasonality in the model and observational data, with total variations of 0.57 and 0.58 Sv, respectively, both with a well-defined maximum in August. When constructing two new time series by removing the seasonality from both observation and model data, the correlation is 0.72. Here the effective number of degrees of freedom[32] is reduced to 39 in comparison with the full sample size of 127 monthly values, corresponding to a decorrelation timescale[32] of 3.2 months. Using standard formulae, we estimate that the probability of getting this correlation of 0.72 for the monthly variability by chance is less than 10^{-7}. Similarly, autocorrelation is insignificant in the annual mean series, which implies that the probability of getting a correlation coefficient as high as 0.92 is less than 10^{-4}. We can therefore state beyond any reasonable doubt that the correlations between model and observations do not occur by chance, implying that the model does indeed represent the dynamic scales responsible for variations in the FBC overflow on monthly, seasonal and interannual scales. Furthermore, modelled variability is not significantly affected by the offset dictated by the climatic mean of unresolved contributions to the total eastern overflow.

A conceptual framework to explain the barotropic/baroclinic pressure compensation. Variations in the simulated FBC overflow transport Δq and pressure difference $\Delta P = \Delta P_{\mathrm{trop}} + \Delta P_{\mathrm{clin}}$ between the relatively quiescent basin downstream of the channel and the Norwegian basin, upstream of the channel, are close to being linearly related at the depth of the overflow (Fig. 3b): $\Delta q = k\Delta P$, where k is about 10^{-3} Sv Pa^{-1}. If sea level in this basin rises by Δh, the pressure difference increases by $\rho g \Delta h$. The associated increase in overflow transport enhances the removal of deep water from the basin, which causes the isopycnal surfaces to deepen, thereby reducing the baroclinic (ΔP_{clin}) component of the pressure difference. To quantify this compensating effect, we approximate the Norwegian basin by a two-layer system with density difference $\Delta\rho$. In this framework, average isopycnal deepening during a period T is equivalent to a deepening (ΔD) of the interface. If A is the area of the interface between the overflow and upper layer, the decrease in the baroclinic pressure difference follows from continuity:

$$\Delta P_{\mathrm{clin}} = -g\Delta\rho\Delta D = \frac{-g\Delta\rho kT}{A}\Delta P_{\mathrm{trop}} \qquad (1)$$

To compensate for the barotropic pressure difference change, the fraction on the right-hand side of (1) should equal unity. This gives a rough estimate of the time T needed for the baroclinic field to respond to an induced change in sea-level difference across the ridge and for the overflow to return to its original value.

Using the area defined by the 500-m isobath (5.8×10^{11} m^2) of the Norwegian basin for the interface area and a typical value of 0.5 kg m^{-3} for $\Delta\rho$, T is about 3 years. This conceptual framework therefore explains the compensation between barotropic and baroclinic contributions to the cross-ridge pressure difference evident in Fig. 3a on these timescales. On longer timescales, the overflow would then be determined by the import of dense water to the Norwegian basin.

30. Kistler, R. E. *et al.* The NCEP–NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bull. Am. Meteorol. Soc.* **82,** 247–268 (2001).
31. Vellinga, M., Dickson, B. & Curry, R. in *Arctic–Subarctic Ocean Fluxes: Defining the Role of the Northern Seas in Climate* (eds Dickson, R. R., Meincke, J. & Rhines, P.) 289–313 (Springer, 2008).
32. von Storch, H. & Zwiers, F. W. *Statistical Analysis in Climate Research* (Cambridge Univ. Press, 1999).

LETTERS

Intraseasonal interaction between the Madden–Julian Oscillation and the North Atlantic Oscillation

Christophe Cassou[1]

Bridging the traditional gap between the spatio-temporal scales of weather and climate is a significant challenge facing the atmospheric community. In particular, progress in both medium-range and seasonal-to-interannual climate prediction relies on our understanding of recurrent weather patterns and the identification of specific causes responsible for their favoured occurrence, persistence or transition. Within this framework, I here present evidence that the main climate intra-seasonal oscillation in the tropics—the Madden–Julian Oscillation[1,2] (MJO)—controls part of the distribution and sequences of the four daily weather regimes defined over the North Atlantic–European region in winter[3]. North Atlantic Oscillation[4] (NAO) regimes are the most affected, allowing for medium-range predictability of their phase far exceeding the limit of around one week that is usually quoted. The tropical–extratropical lagged relationship is asymmetrical. Positive NAO events mostly respond to a mid-latitude low-frequency wave train initiated by the MJO in the western–central tropical Pacific and propagating eastwards. Precursors for negative NAO events are found in the eastern tropical Pacific–western Atlantic, leading to changes along the North Atlantic storm track. Wave-breaking diagnostics tend to support the MJO preconditioning and the role of transient eddies in setting the phase of the NAO. I present a simple statistical model to quantitatively assess the potential predictability of the daily NAO index or the sign of the NAO regimes when they occur. Forecasts are successful in ~70 per cent of the cases based on the knowledge of the previous ~12-day MJO phase used as a predictor. This promising skill could be of importance considering the tight link[4] between weather regimes and both mean conditions and the chances of extreme events occurring over Europe. These findings are useful for further stressing the need to better simulate and forecast the tropical coupled ocean–atmosphere dynamics, which is a source of medium-to-long range predictability and is the Achilles' heel of the current seamless prediction suites[5–7].

Travelling synoptic pressure systems, or storms, contribute to a significant fraction of the daily to interannual variability of the extratropical climate. They are associated with the unstable nature of the upper-level westerly jet stream, and feed circulation patterns of larger scale (weather regimes) in which they are embedded. These regimes can be interpreted as quasi-stationary atmospheric circulations[8] during which the character of the synoptic storms is unusually persistent[9]. They are spatially well defined (typically the width of the oceanic basin) and limited in number; they ideally correspond to statistical-dynamical equilibria in phase space[10]. Objective analyses based on clustering techniques are often used to identify number, spatial structure and frequency of occurrence of weather regimes. The partition algorithm[11] I apply to daily maps of geopotential height (Supplementary Information) leads to four circulation patterns estimated from extended boreal winter days (November–March) over 1974–2007 and over a broad North Atlantic–European domain (Fig. 1).

In agreement with previous studies[3,4], the two first regimes can be viewed as the negative and positive phases of the NAO (NAO− and NAO+, respectively), which can essentially be considered a measure of the variability of the zonal flow over the North Atlantic. The third regime is named Atlantic ridge and the fourth is often referred to as Scandinavian blocking (SBL). Links between flow regimes and mean conditions over Europe have been documented from daily[12,13] to decadal timescales[14]. The chances of extreme events (cold outbreak, heavy rainfall) occurring are also related to the four weather patterns (Supplementary Fig. 2), suggesting that a large part of the statistical distribution for surface variables, even regionally, could be assessed through the weather regime model.

A similar approach in weather-type classes has recently been applied in the tropics to describe and monitor in real time the dominant mode of intraseasonal climate variability, the MJO[1]. This is a natural component of the tropical coupled ocean–atmosphere system, and is characterized by a planetary-scale alternation of wet and dry periods associated with several changes in both tropical and subtropical atmospheric dynamics[2]. The MJO packet propagates eastwards around the globe with a typical 30–70-day cycle. By combining real-time satellite outgoing long-wave radiation (OLR, used as a proxy for convection) and atmospheric dynamical fields from operational reanalysis, it is possible to partition the daily MJO activity into eight intrinsic phases of 7–8-day nominal persistence[15] (Supplementary Information). These phases or classes can be interpreted as the tropical analogues of the extratropical weather regimes, except that regimes have episodic behaviour due to dominant chaos at mid latitudes whereas the time evolution of the MJO phases is mostly oscillatory.

MJO wintertime composites determined from November–March daily maps over 1974–2007 as for weather regimes show (Fig. 2) that the eastwards displacement of the anomalous OLR pattern is associated with significant modifications of the upper-level atmospheric circulation, interpreted as an equatorial Rossby–Kelvin wave response to anomalous tropical heating[16]. Those changes are characterized by vorticity couplets straddling the Equator and either collocated or shifted slightly westwards with respect to the anomalous convection cores and associated large-scale ascent or descent (Supplementary Fig. 3). Note that although OLR signals become weak when the MJO moves to the Western Hemisphere, these anomalous vorticity and divergence patterns clearly remain significant (for example, see Fig. 2, phases 2–3 and 6–7) and truly make manifest the planetary nature of the tropical intraseasonal oscillation.

[1]CNRS-Cerfacs, Global Change and Climate Modelling project, 42 Avenue G. Coriolis, 31057 Toulouse, France.

Extratropical responses to MJO kicks have been extensively described[16–19]. In this study, I adopt a novel approach more relevant to forecasting issues, and combine both tropical and extratropical clusters to investigate how the MJO influences the known and independent North Atlantic modes on medium-range timescales. For the eight MJO phases and for lags of up to +15 days (MJO in advance), I count the number of occurrences of each weather regime and compare this number to its mean (Fig. 3). The relationship between phase 1 of the MJO and the occurrence of the four weather regimes is marginally significant. By contrast, the other MJO phases strongly suggest a significant tropical forcing upon the North Atlantic dynamics. For instance, phase 3 is not discriminative for the NAO

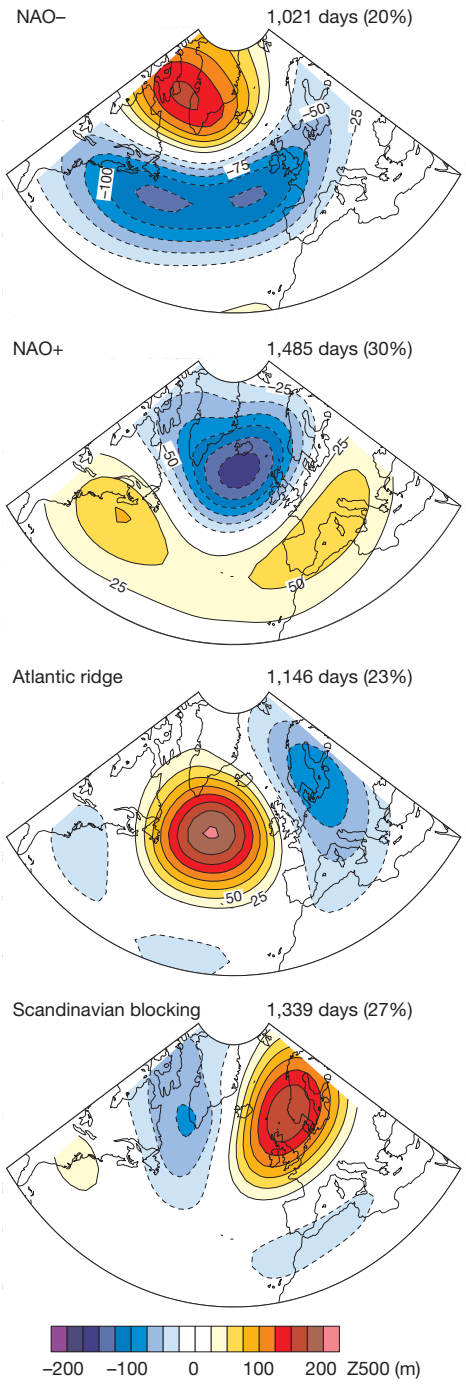


Figure 1 | Wintertime North Atlantic weather regimes. Centroids of the four weather regimes obtained from daily anomalous geopotential height at the 500-hPa altitude (Z500, colour) from the National Center for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) Reanalysis. Each percentage corresponds to the stated number of days and represents the mean frequency of occurrence of the regime computed over 1974–2007 from 1 November to 31 March. Contour intervals are 25 m. Details on the algorithm used for clustering are given in the Methods Summary and Supplementary Information.

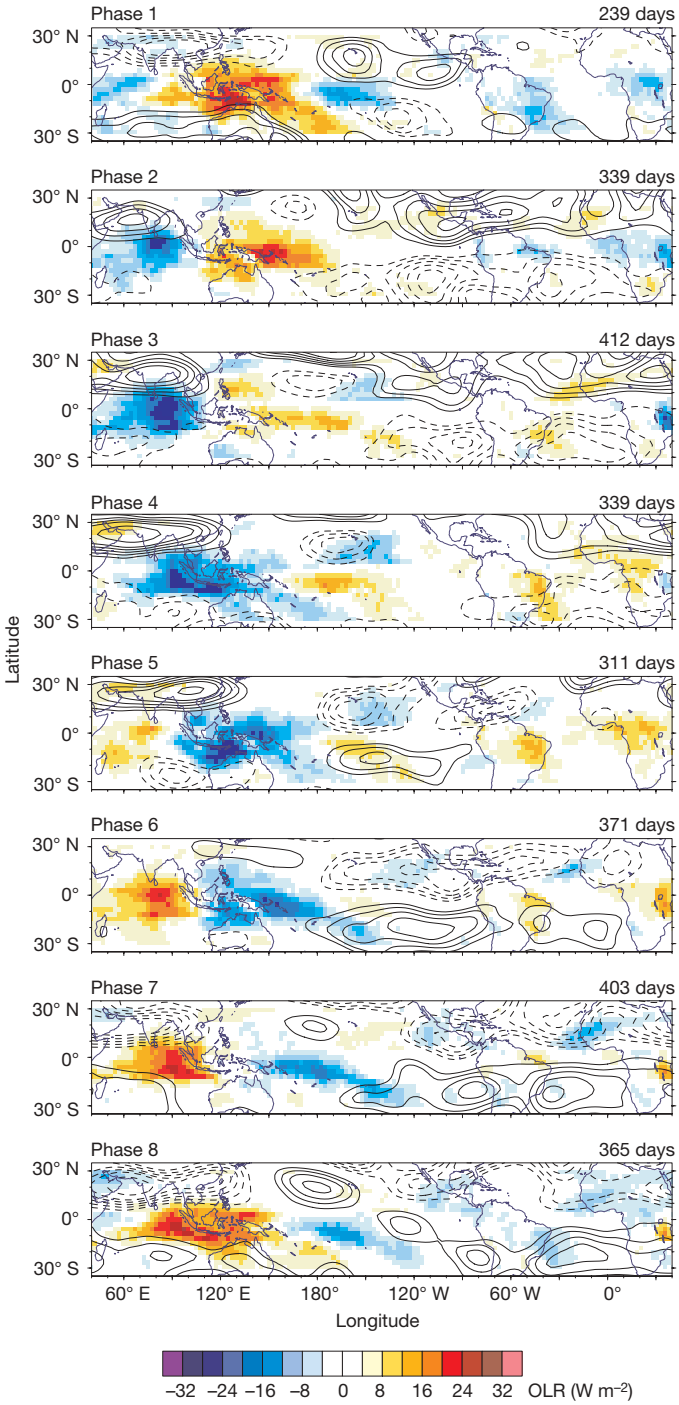


Figure 2 | Dynamical and thermodynamical signatures of the eight phases of the MJO. Wintertime composite of OLR (colour) and stream function anomalies at 300 hPa (PSI300, contours) for the eight phases. Only strong MJO cycles are retained, corresponding to the stated number of days for each phase (see Supplementary Information). Bluer colours correspond to enhanced convection activity and wetter conditions, and redder colours correspond to reduced convection activity and drier conditions. Shading intervals are 4 W m^{-2} for OLR. Contour intervals are 1×10^6 m^2 s^{-1} for PSI300, starting at $\pm 2 \times 10^6$ m^2 s^{-1}. Positive values (solid) in the Northern Hemisphere and negative values (dashed) in the Southern Hemisphere represent anomalous anticyclonic circulation.

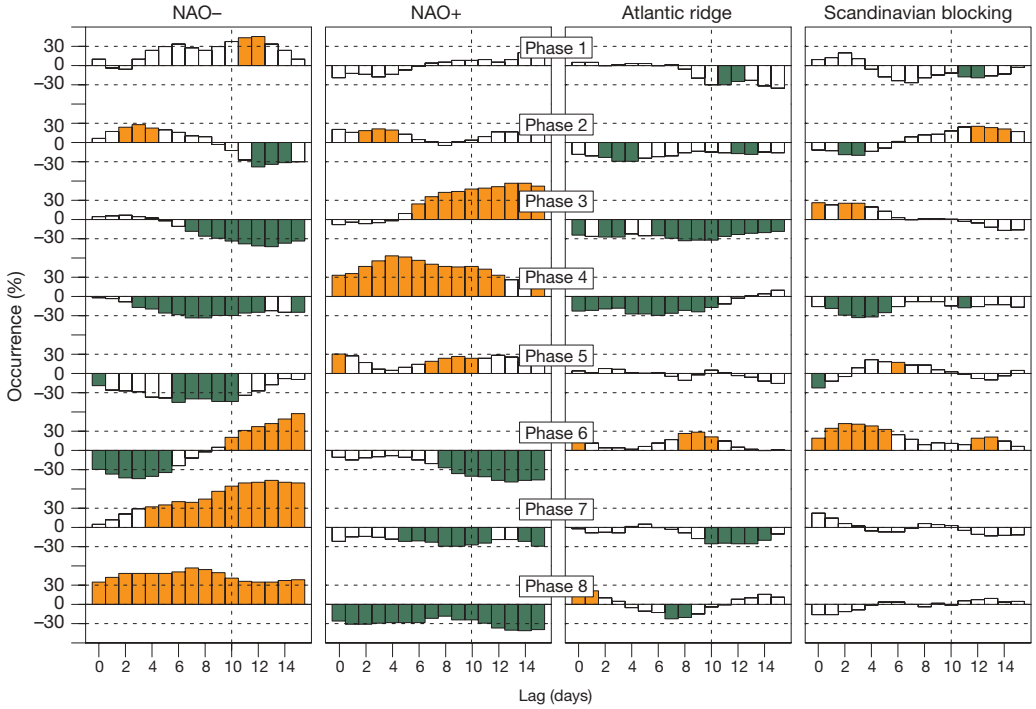


Figure 3 | Lagged relationships between the eight phases of the MJO and the four North Atlantic weather regimes. Table of contingency between the MJO phases (rows) and the North Atlantic weather regimes (columns). For each MJO phase, I plot the anomalous percentage occurrence of a given regime as a function of lag in days (with regimes lagging MJO phases). The 0% value means that the MJO phase is not discriminative for the regime whose occurrence is climatological. A 100% value would mean that this regime occurs twice as frequently as its climatological mean; −100% means no occurrence of this regime. The presence of a slope as a function of lag is suggestive of the MJO forcing. For white bars, either the change in the distribution between the four regimes is not significant on the basis of χ^2 statistics at the 99% significance level, or the individual anomalous frequency of occurrence is lower than the minimum significant threshold tested at 95% using a Gaussian distribution (approximation for binomial distribution because of the sufficiently large sampling). For orange and green bars, the regimes occur significantly more or, respectively, less frequently than their climatological occurrences.

regimes at lag 0, whereas ~10 days later the probability of a NAO− event occurring is reduced by ~40% and is mostly compensated for by an increase of ~60% in the probability of NAO+. The opposite is found for phase 6, which shows an increase of NAO− occurrence probability building up to ~70% for lags greater than 10 days, with NAO+ less probable. The Atlantic ridge regime seems to be less affected by the progress of the MJO, its occurrence being simply reduced by construction when NAO+ or NAO− regimes are dominant (phases 3–4 and 7, respectively). SBL occurrence is also weakly altered, except at short lag time in phase 6 (enhanced excitation) before NAO− maturation. The changes in the regime distribution occur progressively in accordance with the nominal 7–8-day persistence of the eight MJO phases.

Thus, Fig. 3 constitutes a contingency table providing evidence that phases 3 and 6 of the MJO can be interpreted as precursors of NAO+ and NAO− regimes, respectively. Lagged composites for these two specific phases reveal a clear asymmetry in the tropical forcing upon the two regimes. The days following phase 3 (Fig. 4a) are dominated by a mid-latitude low-frequency anomalous wave train that originates in the eastern Pacific, stretches across the North American continent and propagates eastwards following the Northern Hemispheric waveguide. Its penetration along the North Atlantic mean storm track (40°–60° N) is associated with dominant anticyclonic synoptic-scale wave breakings (AWBs) known as precursors for NAO+[20,21] from a lag of +6 days onwards (Fig. 4c), which is consistent with the response to the local modified background flow and the upstream shift towards the Equator of the tail end of the Pacific storm track[21,22] (Supplementary Fig. 5). The opposite picture emerges in the days following phase 6 of the MJO. There is no signal coming from the Pacific (Fig. 4b); height anomalies originating from Europe and propagating westwards are found instead from a lag of 0 to +6 days, before the development of a quasi-standing pattern projecting on NAO−. The proportion of AWBs is then clearly reduced from a lag of +6 days onwards (Fig. 4d). This reduction is almost entirely controlled by very high-frequency transients, which is consistent with the preferred *in situ* development of NAO− events[23] associated with more frequent cyclonic wave breakings[20,21] (CWBs), and contrasts with NAO+ events, in which intermediate-frequency eddy activity has a role[23] (difference between the purple and green curves in Fig. 4c).

The precursory wave train for NAO+ is initiated in the western–central Pacific (phases 2–3) in response to the direct MJO forcing following forced Rossby wave theories (Supplementary Fig. 6). Processes for teleconnection between MJO and NAO− regimes appear to be less straightforward. Two non-exclusive mechanisms are proposed. SBL regimes are present at short lag time in phase 6 (Fig. 3) and may subsequently trigger the onset of NAO− events[24,25]. This hypothesis is consistent with the retrograde propagation of height anomalies[23] (Fig. 4b). The enhanced occurrence of SBL in phase 6, and consequently NAO− at greater lags, can be interpreted as the consequence of previous NAO+ excitation (phases 4–5), considering that the preferred transitions between regimes follow the route NAO+ to SBL to NAO−[3] (Supplementary Information). The hypothesis is based on evidence that the extratropical atmosphere retains information about regime occurrences over medium-range timescales[3]. In this case, enhanced NAO− in late phase 6 and phase 7 would not be directly forced by the MJO, but would correspond to the timescale resonance between the eastward propagation of the MJO and the preferred sequence of the North Atlantic regimes.

The second mechanism proposed for teleconnection between MJO and NAO− relies on direct tropical forcings originating from the

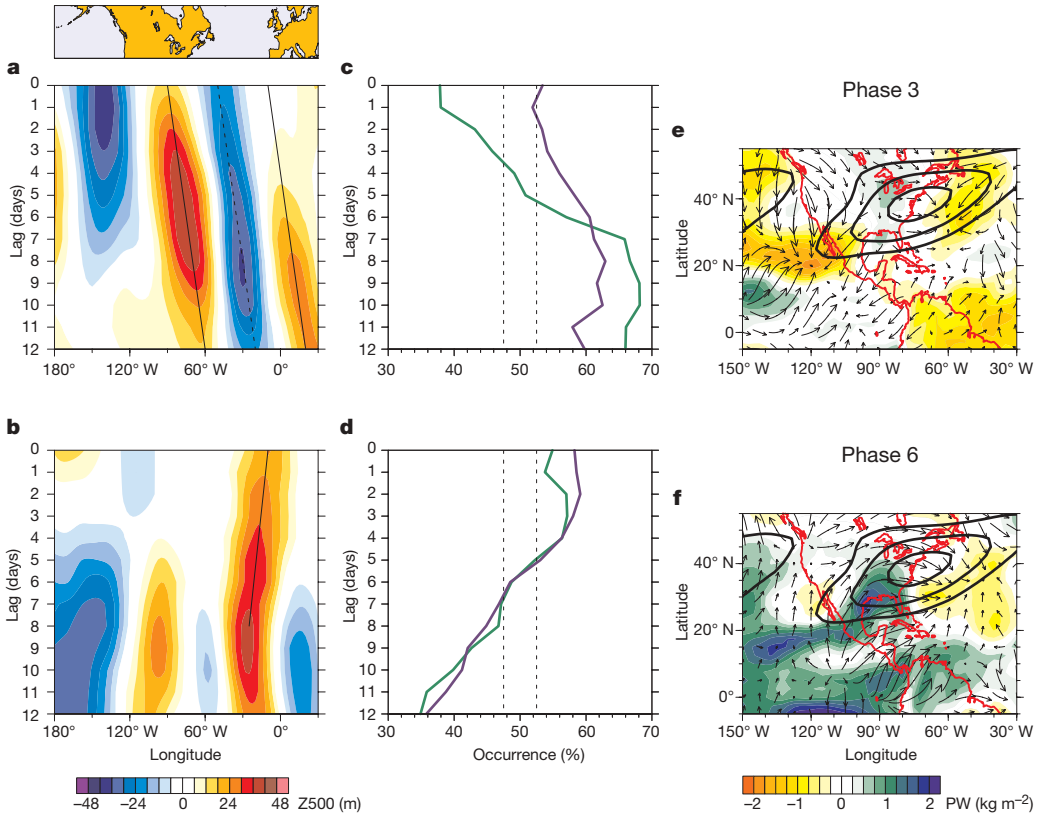


Figure 4 | Asymmetrical tropical–extratropical connection between two specific phases of the MJO leading to NAO+ and NAO− events. **a**, **b**, Longitude–time plots of lagged composites of daily Z500 (colour) anomalies averaged between 40° N and 60° N for phases 3 and 6 of the MJO, respectively. **c**, **d**, Percentage occurrence of AWB as a function of lag for phases 3 and 6, respectively. Wave breakings are assessed from a daily index of meridional eddy momentum flux, band-pass filtered either for 2–6 days (very high frequency, purple) or for 2–12 days (intermediate plus very high frequency, green), and averaged along the Atlantic storm track (100°–10° W, 30°–60° N). Positive and negative values of this index respectively stand for AWB and CWB[22]. The percentage occurrence of AWB is thus computed as the ratio of AWB days to the total sample. Significance at 95% is tested using a Monte Carlo method (1,000 resamplings) and indicated by the two dashed lines. **e**, **f**, Composites of anomalous 300-hPa divergent wind (arrows) and precipitable water (PW, colour) averaged from lags from 0 to +5 days, for phases 3 and 6, respectively. Shading intervals are 0.25 kg m^{-2} for precipitable water. The contours represent the mean climatological jet stream (zonal wind at 300 hPa). Contour intervals are 6 m s^{-1}, starting at 22 m s^{-1}.

eastern Pacific. At short lag time, although the anomalous convection is weak (Fig. 2), phase 6 is associated with tropical upper-level divergence around 120° W (the signal is just as strong as for phase 3 over Indonesia, as indicated in Supplementary Fig. 3) and is compensated to the northeast by convergence both at the entrance to, and the southern flank of, the mean climatological jet stream (Fig. 4f). The local Hadley cell is intensified[26], leading to advection of absolute vorticity by the MJO divergent tropical outflow (Fig. 4f) and to enhanced momentum convergence around 30° N. This picture is consistent with there being a Rossby wave source[27] around 20° N, 110° W (Supplementary Fig. 6) that initiates a downstream wave train propagating northeastwards towards Europe following a preferred curving path (not shown) in line with several studies[16,28]. Note that, concurrently, enhanced precipitable water extends from the tropical eastern Pacific to the entrance of the jet stream during phase 6, consistent with the upper-tropospheric anomalous circulation. There is evidence[22] that destabilized atmosphere due to enhanced moisture located upstream from the North Atlantic storm track favours CWBs in agreement with our findings, and could thus be an additional contributor to NAO− occurrence. Opposite phenomena (Fig. 4e) tend to appear for phase 3 (upper-level convergence at ~110° W, tendency towards drier atmosphere at the entrance of the jet stream), although signals are significantly weaker.

In summary, I show that forced Rossby waves initiated by the MJO either in the western Pacific (NAO+ precursors) or in the eastern Pacific (NAO− precursors) should be interpreted as a modification of the background flow[29] or catalysts for a full development of NAO regimes that involves a strong interaction with transient eddies by means of wave breaking.

These results contribute to the window of opportunity for enhanced predictive skill of the wintertime NAO at medium-range timescales. As a crude estimate, I built the most simple generalized linear model[30] using MJO phases as predictors (Supplementary Information). The probability of predicting the correct sign of the NAO regimes when they occur is ~70% at a lead time of +9 to +13 days (Supplementary Figs 7–8). A similar accuracy is obtained using the NAO index subdivided into terciles. These significant values should be treated as benchmarks for numerical weather prediction models and mean that there is some predictability far exceeding the limit of around one week that is usually quoted as the limit for the North Atlantic–European sector. On the basis of my findings, I suggest that the next step towards improving our capabilities in forecasting on the timescale of 5–30 days can be summarized in three main points. First, it is essential to get very accurate real-time information on the three dimensional tropical atmospheric dynamics and thermodynamics for the model initialization step. Second, it is crucial to be able to correctly simulate and forecast the evolution or persistence of the initialized MJO dynamics that is somehow achieved using statistical models[7] but which is far from being possible with the current generation of weather[5] and climate models[6,7]. Last, it is important to correctly simulate the tropical–extratropical atmospheric bridges from daily to seasonal timescales. This requires a good representation of the atmospheric mean state and transient eddy activity, upon which the latitudinal connection is dependent.

METHODS SUMMARY

North Atlantic weather regimes. The *k*-means clustering algorithm is applied to anomalous daily geopotential height at 500 hPa from the NCEP/NCAR Reanalysis to obtain four weather regimes over a large North Atlantic–Europe domain. The decomposition is limited to the 90° W–30° E, 20°–80° N region and performed in the empirical-orthogonal-function phase space to speed up the computation (14 modes retained, corresponding to 90% of variance). A cosine weight as a function of latitude is applied to the data and the climatology is computed over 1974–2007 for winter days (November–March). Daily attribution to a given weather pattern is based on a minimum-Euclidean-distance criterion. The optimal number of regimes, *k*, is chosen on the basis of Brownian noise statistics[11].

MJO. The MJO activity is assessed through the real-time multivariate MJO index (RMMI)[15] provided daily and operationally by the Australian Bureau of Meteorology. Only active MJO phases (RMMI amplitude greater than one) are retained throughout the entire study.

MJO–North Atlantic coupling. Statistical evidence of interaction between the MJO and North Atlantic regimes is simply assessed through lagged composites between the MJO phases (in advance) and the North Atlantic regimes. Significance for contingency is based on χ^2 and Gaussian statistics to test respectively for changes in the combined four-regime distribution and for individual anomalous occurrences as a function of lag in days. The physical mechanisms supporting the asymmetrical MJO–NAO connection are extracted from lagged composites for several dynamical fields from the NCEP/NCAR Reanalysis. Diagnostics for wave breaking are obtained from meridional eddy momentum fluxes computed daily over the North Atlantic along the storm track, and analysed in two bands of frequency based on Lanczos filtering.

Received 29 April; accepted 9 July 2008.

1. Madden, R. A. & Julian, P. R. Observations of the 40–50 day tropical oscillation. *Mon. Weath. Rev.* **112,** 1109–1123 (1994).
2. Zhang, C. Madden-Julian oscillation. *Rev. Geophys.* **43,** doi:10.1029/2004RG000158 (2005).
3. Vautard, R. Multiple weather regimes over the North Atlantic: Analysis of precursors and successors. *Mon. Weath. Rev.* **118,** 2056–2081 (1990).
4. Hurrell, J. W., Kushnir, Y., Ottersen, G. & Visbeck, M. in *North Atlantic Oscillation: Climate Significance and Environmental Impact* (eds Hurrell, J. W., Kushnir, Y., Ottersen, G. & Visbeck, M.) 1–35 (Geophys. Monogr. 134, American Geophysical Union, 2003).
5. Vitard, F., Woolnough, S., Balmaseda, M. A. & Tompkins, A. M. Monthly forecast of the Madden-Julian Oscillation using a coupled GCM. *Mon. Weath. Rev.* **135,** 2700–2715 (2007).
6. Lin, J.-L. *et al.* Tropical intraseasonal variability in 14 IPCC AR4 climate models. Part I: convective signals. *J. Clim.* **19,** 2665–2690 (2006).
7. Waliser, D. *et al.* The experimental MJO prediction project. *Bull. Am. Meteorol. Soc.* **87,** 425–431 (2006).
8. Reinhold, B. & Pierrehumbert, R. Dynamics of weather regimes: quasi-stationary waves and blocking. *Mon. Weath. Rev.* **110,** 1105–1145 (1982).
9. Straus, D., Corti, S. & Molteni, F. Circulation regimes: chaotic variability versus SST-forced predictability. *J. Clim.* **20,** 2251–2272 (2007).
10. Molteni, F., Kuscharski, F. & Corti, S. in *Predictability of Weather and Climate* (eds Palmer, T. & Hagedorn, R.) 365–389 (Cambridge Univ. Press, 2006).
11. Michelangeli, P., Vautard, R. & Legras, B. Weather regimes: recurrence and quasi-stationarity. *J. Atmos. Sci.* **52,** 1237–1256 (1995).
12. Philipp, A. *et al.* Long term variability of daily North Atlantic-European pressure patterns since 1850 classified by simulated annealing clustering. *J. Clim.* **20,** 4065–4095 (2007).
13. Slonosky, V. C. & Yiou, P. The North Atlantic Oscillation and its relationship with near surface temperature. *Geophys. Res. Lett.* **28,** 807–810 (2001).
14. Hurrell, J. W. Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. *Science* **26,** 676–679 (1995).
15. Wheeler, M. C. & Hendon, H. H. An all-season real-time multivariate MJO index: development of an index for monitoring and prediction. *Mon. Weath. Rev.* **132,** 1917–1932 (2004).
16. Matthews, A. J., Hoskins, B. J. & Masutani, M. The global response to tropical heating in the Madden-Julian Oscillation during northern winter. *Q. J. R. Meteorol. Soc.* **130,** 1991–2011 (2004).
17. Ferranti, L., Palmer, T. N., Molteni, F. & Klinker, E. Tropical-extratropical interaction associated with the 30–60 day oscillation and its impact on medium and extended range prediction. *J. Atmos. Sci.* **47,** 2177–2199 (1990).
18. Higgins, R. W. & Mo, K. C. Persistent North Pacific circulation anomalies and the tropical intraseasonal oscillation. *J. Clim.* **10,** 223–244 (1997).
19. Zhou, S. & Miller, A. J. The interaction of the Madden-Julian Oscillation and the Arctic Oscillation. *J. Clim.* **18,** 143–159 (2005).
20. Benedict, J. J., Lee, S. & Feldstein, S. B. Synoptic view of the North Atlantic Oscillation. *J. Atmos. Sci.* **61,** 121–144 (2004).
21. Franske, C., Lee, S. & Feldstein, S. B. Is the North Atlantic Oscillation a breaking wave? *J. Atmos. Sci.* **61,** 145–160 (2004).
22. Rivière, G. & Orlanski, I. Characteristics of the Atlantic storm track eddy activity and its relationship with the North Atlantic Oscillation. *J. Atmos. Sci.* **64,** 241–266 (2007).
23. Feldstein, S. B. The dynamics of NAO teleconnection pattern growth and decay. *Q. J. R. Meteorol. Soc.* **129,** 901–924 (2003).
24. Croci-Maspoli, M., Schwierz, C. & Davies, H. Atmospheric blocking: space-time links to the NAO and PNA. *Clim. Dyn.* **29,** 713–725 (2007).
25. Scherrer, S. C., Croci-Maspoli, M., Schwierz, C. & Appenzeller, C. Two dimensional indices of atmospheric blocking and their relationship with winter climate patterns in the Euro-Atlantic region. *Int. J. Climatol.* **26,** 233–249 (2006).
26. Tyrrell, G. C., Karoly, D. J. & McBride, J. L. Links between tropical convection and variations of the extratropical circulation during TOGA-CORE. *J. Atmos. Sci.* **53,** 2735–2748 (1996).
27. Sardeshmukh, P. & Hoskins, B. The generation of global rotational flow by steady idealized tropical divergence. *J. Atmos. Sci.* **45,** 1228–1251 (1988).
28. Hoskins, B. J. & Ambrizzi, T. Rossby wave propagation on a realistic longitudinally varying flow. *J. Atmos. Sci.* **50,** 1661–1671 (1993).
29. Woollings, T., Hoskins, B., Blackburn, M. & Berrisford, P. A new Rossby wave-breaking interpretation of the North Atlantic Oscillation. *J. Atmos. Sci.* **65,** 609–626 (2008).
30. Simon, S. J. & Baddour, O. in *Seasonal Climate: Forecasting and Managing Risk* (eds Troccoli, A., Harrison, M., Anderson, D. L. T. & Mason, S. J.) 163–201 (Springer, 2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The author wishes to thank L. Terray, G. Rivière, R. Madden, G. Madec and C. Périgaud for discussions. The author also thanks F. Chauvin, H. Douville and S. Valcke for comments on the manuscript. The author is very grateful to O. Mestre and J.-P. Céron for their help with statistics. The figures were produced using the NCL software developed at NCAR. This work was supported by CNRS and by the European Union's Sixth Framework Programme (DYNAMITE and ENSEMBLES).

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to C.C. (cassou@cerfacs.fr).

LETTERS

Earliest date for milk use in the Near East and southeastern Europe linked to cattle herding

Richard P. Evershed[1], Sebastian Payne[2], Andrew G. Sherratt[3]‡, Mark S. Copley[1], Jennifer Coolidge[4], Duska Urem-Kotsu[5], Kostas Kotsakis[5], Mehmet Özdoğan[6], Aslý E. Özdoğan[7], Olivier Nieuwenhuyse[8], Peter M. M. G. Akkermans[8], Douglass Bailey[9], Radian-Romus Andeescu[10], Stuart Campbell[11], Shahina Farid[12], Ian Hodder[13], Nurcan Yalman[14], Mihriban Özbaşaran[6], Erhan Bıçakcı[6], Yossef Garfinkel[14], Thomas Levy[15] & Margie M. Burton[15]

The domestication of cattle, sheep and goats had already taken place in the Near East by the eighth millennium BC[1–3]. Although there would have been considerable economic and nutritional gains from using these animals for their milk and other products from living animals—that is, traction and wool—the first clear evidence for these appears much later, from the late fifth and fourth millennia BC[4,5]. Hence, the timing and region in which milking was first practised remain unknown. Organic residues preserved in archaeological pottery[6,7] have provided direct evidence for the use of milk in the fourth millennium in Britain[7–9], and in the sixth millennium in eastern Europe[10], based on the $\delta^{13}C$ values of the major fatty acids of milk fat[6,7]. Here we apply this approach to more than 2,200 pottery vessels from sites in the Near East and southeastern Europe dating from the fifth to the seventh millennia BC. We show that milk was in use by the seventh millennium; this is the earliest direct evidence to date. Milking was particularly important in northwestern Anatolia, pointing to regional differences linked with conditions more favourable to cattle compared to other regions, where sheep and goats were relatively common and milk use less important. The latter is supported by correlations between the fat type and animal bone evidence.

The use of milk, wool and traction, so-called 'secondary' products, obtained from domestic animals without killing them, marks an important step in the history of domestication[4,5]. But evidence for when and how this first happened is inconclusive. Some researchers have argued that once animals were domesticated the potential benefits of these products would have been exploited rapidly[11]. Others have pointed to the late appearance of unequivocal evidence—that is, representations of milking scenes, carts and ploughs—and to barriers, such as lactose intolerance in humans, suggesting that early domestication was predominantly for meat and hides, postulating a 'secondary products revolution' during the fifth or fourth millennium BC, 2,000–4,000 years after the first domestication of cattle, sheep and goats in the Near East and Europe[5,12]. Evidence provided by figurines and pictures of animals before 4000 BC, and from artefacts (for example, ceramic strainers), has been variously interpreted[13], as has evidence from animal bone assemblages, especially the ages at which animals were killed, taken as reflecting what they were kept for and how they were managed[14–16].

The analysis of lipid residues from pottery, particularly our discovery that ruminant milk fatty acids can be distinguished from those of carcass fats, provided a new tool for detecting early milk use[6,7]. The approach rests upon differences in the $\delta^{13}C$ value of the $C_{18:0}$ (in $C_{x:y}$, x is the number of carbon atoms in the fatty acid, and y is the number of double bonds) fatty acid of milk and carcass fats. This arises from a greater proportion of dietary carbohydrate-derived carbon being used in the biosynthesis of carcass fat $C_{18:0}$, compared to milk fat, up to 40% of which derives from biohydrogenated dietary unsaturated C_{18} fatty acids ($C_{18:3}$, $C_{18:2}$ and $C_{18:1}$)[17,18]. Using this approach, we recently provided evidence for widespread milk use at some of the earliest Neolithic sites in southern Britain[7–9]. However, these sites, dating to the early fourth millennium BC, are late in relation to the Neolithic and Chalcolithic of the Near East and southern and central Europe. The same technique has also provided evidence for milk use in Romania before 5000 BC[10].

Reported here are results from analyses of organic residues from sherds of pottery vessels from fifth- to seventh-millennium BC sites in southeastern Europe, Anatolia and the Levant. Vessels most likely to have been used for food preparation were selected to test where milk use started, and whether the use of milk products first began in the region where farming was pioneered, namely within the Fertile Crescent, or whether it was an innovation of other regions. Figure 1 shows the locations of the 23 sites from which the sherds were sampled. The results of the analyses of 2,225 sherds are summarized in Table 1 and Figs 2 and 3; 12% of the sherds (255) yielded sufficient residue for compound-specific stable carbon isotope analysis. Typical gas chromatographic profiles of the residues displayed in Fig. 2 show that the $C_{16:0}$ and $C_{18:0}$ fatty acids predominate, the high abundance of the latter confirming that the residues derive from animal fats. Mean lipid concentrations varied over the range 0.54–1.74 mg per g sherd. The lower concentrations and incidences of lipid residues in these assemblages, compared to pottery from northern European sites, probably relates to differences in vessel use, clay type, the greater age of the pottery and/or degradative factors associated

[1]Organic Geochemistry Unit, Bristol Biogeochemistry Research Centre, School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK. [2]English Heritage, 1 Waterhouse Square, 138–142 Holborn, London EC1N 2ST, UK. [3]Department of Archaeology, University of Sheffield. [4]Research Laboratory for Archaeology and the History of Art, University of Oxford, 6 Keble Road, Oxford OX1 3QJ, UK. [5]Department of Archaeology, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece. [6]Prehistory Department, Istanbul University, Istanbul 34134, Turkey. [7]Archaeology Department, Çanakkale Onsekiz Mart Üniversitesi, Çanakkale 17020, Turkey. [8]Netherlands National Museum of Antiquities and Leiden University, PO Box 1114, 2301 EC Leiden, The Netherlands. [9]School of History and Archaeology, Humanities Building, University of Cardiff, Colum Drive, Cardiff CF10 3EU, UK. [10]Romanian National Museum of History, Calea Vitoriei, nr. 12, Sect. 3, cod poştal 030026, Bucureşti, Romania. [11]School of Arts, Histories and Cultures, University of Manchester, Oxford Road, Manchester M13 9PL, UK. [12]Institute of Archaeology, University College, London, 31–34 Gordon Square, London WC1H 0PY, UK. [13]Archaeology Center, Stanford University, Stanford, California 94305, USA. [14]Institute of Archaeology, Hebrew University of Jerusalem, Jerusalem 91905, Israel. [15]Department of Anthropology, University of California San Diego, 9500 Gilman Drive, La Jolla, California 92093-0532, USA.
‡Deceased.

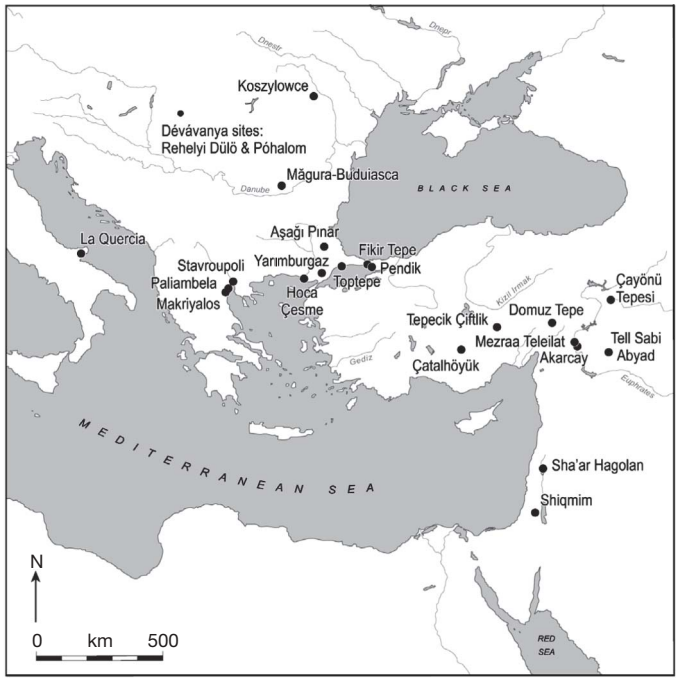


Figure 1 | Map showing the locations of sites providing pottery for organic residue analysis.

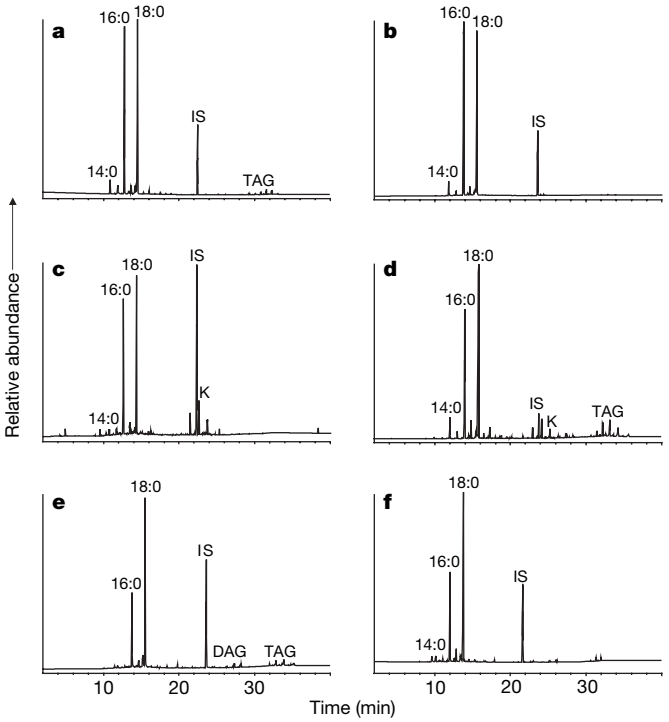


Figure 2 | Partial gas chromatograms of total lipid extracts from pottery. Pottery was from: **a**, Mağura (southeastern Europe); **b**, Makriyalos (northern Greece); **c**, Pendik (northwestern Anatolia); **d**, Çatalhöyük (central Anatolia); **e**, Cayönü Tepesi (southeastern Anatolia); and **f**, Tell Sabi Abyad (Levant). Abbreviations: N:0, fatty acids with N carbon atoms and no double bonds; K, mid-chain ketones with 31, 33 and 35 carbon atoms; DAG, diacylglycerols; TAG, triacylglycerols; and IS, internal standard (*n*-tetratriacontane). Sample reference numbers: **a**, MAG25; **b**, MAK100; **c**, PEN7; **d**, CAT180; **e**, CT53; **f**, SAB21.

with discard and burial. Pyrolytically formed long-chain ketones were present in a number of the vessels, consistent with them having being extensively heated during use[19]. All extracts were investigated for components likely to produce false positives for dairy fats. For example, sherds yielding wax esters diagnostic of beeswax were not included in the stable isotopic analyses[20]. The very low abundance or absence of detectable long-chain *n*-alkanoic (that is, C_{20} and C_{22}), isoprenoid and ω-(*o*-alkylphenyl)-
alkanoic acids rules out significant contributions from aquatic resources[21].

Compound-specific stable carbon isotope analyses yielded $\delta^{13}C$ values for $C_{16:0}$ fatty acids in the range −30‰ to −21‰. This range is somewhat wider than that seen for northern European sites (−30‰ to −25‰) as a result of the contribution of C_4 (ref. 22) and water-stressed[23] plants to the diets of domesticated animals in parts of the region. By plotting (Fig. 3) $\Delta^{13}C$ ($= \delta^{13}C_{18:0} - \delta^{13}C_{16:0}$)

Table 1 | Details of sites, dates, sherds, and lipids and their concentrations

Region	Site	Date (kyr BC)	Number of sherds		Lipid classes detected	Lipid concentration max/mean ($mg\,g^{-1}$)
			Total analysed	With $>5\,\mu g\,g^{-1}$ lipid		
Central/	Koszylowce	4.5–3.5*	339	22	FFA, TAG, WE	0.90/0.08
SE Europe	Póhalom	4.5–4.0†				
	La Quercia	5.5–4.5*				
	Măgura	5.5–5.2†				
	Rehelyi Dülö	6.0–5.5†				
N Greece	Makriyalos	5.2–4.9*	305	56	FFA, K, WE, TAG	1.74/0.06
	Stavroupoli	5.7–4.2†				
	Paliambela	6.0–4.2*				
NW Anatolia	Aşağı Pınar	5.5–5.0‡	703	102	FFA, K, WE, TAG	0.06/0.06
	Toptepe	5.5–5.0‡				
	Yarımburgaz	6.0–5.5‡				
	Fikir Tepe	6.0–5.5‡				
	Hoca Çesme	6.5–5.5‡				
	Pendik	6.5–6.0‡				
Central Anatolia	Domuztepe	5.9–5.5*	187	34	FFA, K, TAG	0.90/0.08
	Tepecik Çiftlik	5.9–5.6*				
	Çatalhöyük	7.0–6.0†				
SE Anatolia	Akarçay Tepe	7.0–6.2*	236	13	FFA, K, WE, TAG	1.63/0.28
	Çayönü Tepesi	6.5–6.0*				
	Mezraa Teleilat	6.5–6.0*				
Levant	Tell Sabi Abyad	6.5–6.0†	448	28	FFA, K, TAG	0.58/0.06
	Shiqmim	4.8–3.5*				
	Sha'ar Hagolan	6.4–6.0*				

FFA, free fatty acids; K, ketones; WE, wax esters (derived from degraded beeswax); TAG, triacylglycerols.
* Milk fats undetectable.
† < 30% milk fats.
‡ > 30% milk fats.

values, any effects on the classification of animal fats are removed, emphasizing biosynthetic and metabolic characteristics of the fat source[7]. These plots remove selective grazing, browsing or foddering differences between cattle and sheep/goats on fat classifications. Figure 3 presents the carbon isotope plots for the four main regions, showing the classifications of the organic residues to animal fat source.

The most striking feature of the data obtained is the emphatic evidence for extensive processing of dairy products in the pottery from all the sites of northwestern Anatolia (Fig. 3a) dating from about 6500–5000 BC, around the Sea of Marmara. Of the ~700 sherds analysed from the six sites considered in this region—Aşağı Pınar (5500–5000 BC), Toptepe (5500–5000), Yarımburgaz (6000–5500), Fikir Tepe (6000–5500), Hoca Çesme (6500–5500) and Pendik (6500–6000)—about 100 (~15%) yielded appreciable animal fat residues, of which 70% contained predominantly dairy fat residues. Thus, the milking of ruminant animals was clearly practised intensively in the sixth and seventh millennia BC in northwestern Anatolia.

Additional support for the latter interpretations comes from correlations with animal bone evidence. A significant feature of the northwestern Anatolian group of sites is that, where data are available, the proportion of cattle bones in animal bone assemblages is considerably higher than in sites in the other areas, presumably reflecting higher rainfall and greener grazing. A positive correlation ($R^2 = 0.56$) also exists between the proportion of sherds with ruminant milk $\Delta^{13}C$ values from different sites, and the relative importance of cattle in animal bone assemblages from the same sites (Fig. 4a), parallelling findings from British Neolithic sites[9]. Thus, the strong evidence for milk use at the northwestern Anatolian sites can reasonably be related to the importance of cattle in the bone assemblages—although ageing data for cattle and sheep/goats from Fikirtepe and Aşağı Pınar are suggestive of mixed use rather than of specialized milk production[24]. Our results suggest that milk was also used in the other areas studied, but was less important. This accords with the available ageing data: for example, sheep and goats at Çatalhöyük were almost all killed as subadults and young adults, a pattern suggestive of concentration on meat production[25]. Significantly, the proportion of sherds with $\Delta^{13}C$ values characteristic of pig fats from different sites is strongly positively correlated (Fig. 4b; $R^2 = 0.85$) with the estimated relative importance of pig in animal bone assemblages from the same sites, analogous to trends observed in late Neolithic sites in northern Europe[26].

Our results provide new insights into the emergence of dairying as a component of the domestication of animals. The appearance of dairy products at early sites in the region is the earliest evidence so far, by ~1–2 millennia, dating back to the start of ceramics in the region; this indicates an earlier date for the milking of domesticated animals than predicted by reconstructions based on other lines of evidence[3]. Significantly, the high incidence of dairy products in pottery from sites in northwestern Anatolia points to intensification of the milking of ruminant animals, at locations remote from the original region of domestication, namely the Fertile Crescent.

Importantly, the results suggest a pattern of regional variation in the importance of milk use rather than of general change with time. Milk appears to have been particularly important in the sites from northwestern Anatolia, ranging in date from the end of the seventh

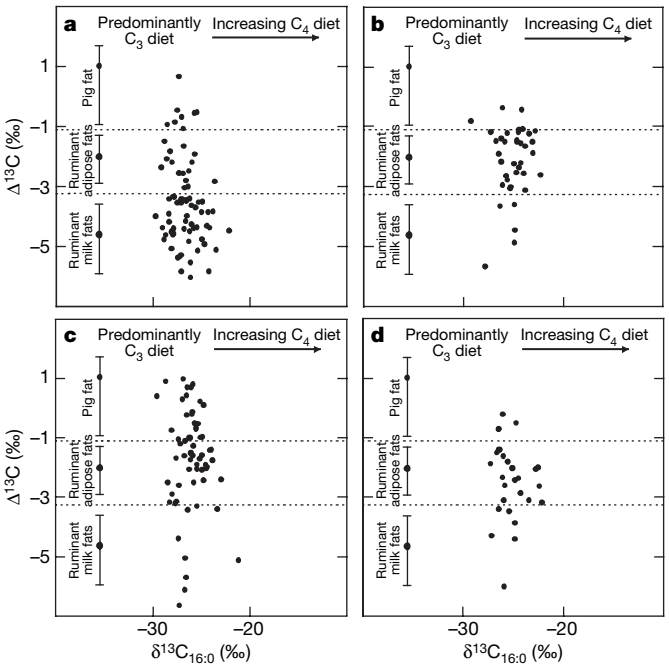


Figure 3 | Plots of the $\Delta^{13}C$ values for archaeological animal fat residues in Neolithic pottery. Pottery was from: **a**, northwestern Anatolia; **b**, central Anatolia; **c**, southeastern Europe/northern Greece; and **d**, eastern Anatolia and the Levant. The $\Delta^{13}C$ values ($= \delta^{13}C_{18:0} - \delta^{13}C_{16:0}$) for the ruminant dairy fats are more depleted than the ruminant adipose fats; the difference in the means is ~2.8‰ which is highly significant (*t*-test; $P < 0.0005$). Pig fats have positive $\Delta^{13}C$ values which do not exhibit significant variance and the differences in the mean values are also highly significant (ANOVA; $P < 0.0005$ between all three commodity groups; Bonferroni adjustment applied). $\delta^{13}C = [(^{13}C/^{12}C)_{sample}/(^{13}C/^{12}C)_{standard}] - 1$, expressed in per mil. All $\delta^{13}C$ values are relative to Vienna PeeDee Belemnite (VPDB) international standard.

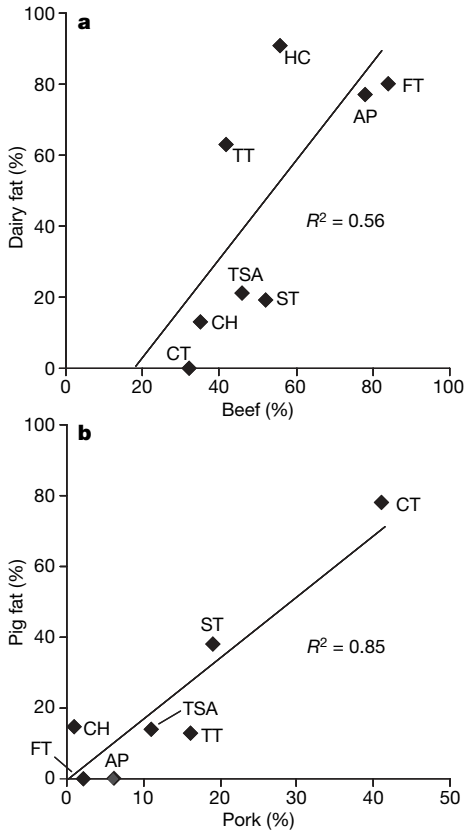


Figure 4 | Percentage animal fat types in pottery versus meat yields based on faunal remains. Both data sets available for: Aşağı Pınar (AP), Çatalhöyük (CH), Çayönü Tepesi (CT), Fikir Tepe (FT), Hoca Çeşme (HC), Tell Sabi Abyad (TSA), Stavroupoli (ST) and Toptepe (TT). Percentage pork and beef is based on ~1,000 to 49,000 identified bones with weightings for pigs (×2) and sheep/goats (×5) to allow for carcass weight and recovery differentials. Percentage dairy and pig fats is based on proportions of 144 residues falling into dairy (**a**) and pig (**b**) fat ranges (Fig. 3). Where more than one bone assemblage is available within periods covered by pottery residues, percentages represent pooled data.

millennium to the fifth millennium BC, contrasting markedly with results from southeastern and central Anatolia at the same time. In neither area is there any strong suggestion of chronological change. In northern Greece, sites dating from the sixth to the fourth millennia BC gave low proportions of sherds with ruminant milk fats; the one site with intermediate results was the sixth millennium Romanian site of Măgura-Buduiasca, which concurs with an earlier study of two sites from this region[10].

Our earlier experiments showed that raw milk lipids absorbed by reproduction ceramics are rapidly destroyed by burial[6,9,27,28], suggesting that the high frequency of ruminant milk lipids from the northwestern Anatolian sites is indicative of milk being processed. Processing milk would have had two important advantages, providing a means of storing surplus milk as products, that is cheese, ghee, and so on, making them available throughout the year, and providing a solution for any problems of lactose intolerance; most lactose intolerant people have fewer problems with consuming processed milk products.

In summary, our findings take the early history of milk use back to the seventh millennium BC, early in the evolution of animal domestication and pottery production and use. The results are significant also for two other reasons: first, they suggest that even at this date (before 6500 BC) milk was processed, making possible the storage of milk products and providing an explanation why, in spite of lactose intolerance, milk use could be adopted quickly, and second, they add to increasing indications of regional differences during the early Neolithic and into the Chalcolithic. Thus, early farming appears not to have been a fixed package; instead, it developed in different ways in different areas, probably in response partly to different environmental conditions and partly to different cultural choices of early farmers.

METHODS SUMMARY

A total of 2,225 well-stratified potsherds were sampled from 23 different archaeological sites across the Near East and southeastern Europe, dated to the Neolithic and Chalcolithic cultural periods. These are grouped within six regions, as follows: central Europe/Ukraine/southeastern Europe (Koszylowce, La Quercia, Măgura, Póhalom and Rehelyi Dülö), northern Greece (Makriyalos, Paliambela and Stavroupoli), northwestern Anatolia (Aşağı Pınar, Fikir Tepe, Hoca Çesme, Pendik, Tepecik Ciftlik, Toptepe and Yarımburgaz), central Anatolia (Çatalhöyük and Domuztepe), southeastern Anatolia (Akarçay, Çayönü Tepesi and Mezraa Teleilat), and the Levant (Tell Sabi Abyad, Sha'ar Hagolan and Shiqmim). Coarsewares and mid-profile sherds were sampled, as these have been found to be the most likely to yield lipid residues[29].

Lipid analyses and interpretations were performed using established protocols described in detail in earlier publications[6,7,9,26,30]. Briefly, ~3 g potsherds were taken and their surfaces cleaned using a modelling drill to remove any exogenous lipids. The sherds were then ground to a powder, an internal standard added and solvent extracted by ultrasonication (chloroform/methanol, 2:1 v/v, 10 ml). The solvent was evaporated under a gentle stream of nitrogen to obtain the total lipid extract (TLE). Aliquots of the TLEs were then trimethylsilylated (*N,O*-bis(trimethylsilyl)trifluoroacetamide 20 µl; 70 °C, 60 min), and submitted to analysis by gas chromatography (GC) and GC/mass spectrometry. Further aliquots of the TLE were treated with $NaOH/H_2O$ (9:1 v/v) in methanol (5% v/v, 70 °C, 1 h). Following neutralization, lipids were extracted into hexane and the excess solvent evaporated under a gentle stream of nitrogen. Fatty acid methyl esters (FAMEs) were prepared by reaction with BF_3-methanol (14% w/v, 70 °C, 1 h). The methyl ester derivatives were extracted with chloroform and the solvent removed under nitrogen. The FAMEs were re-dissolved into hexane for analysis by GC and GC-combustion-isotope ratio MS.

Received 19 February; accepted 19 June 2008.
Published online 6 August 2008.

1. Clutton-Brock, J. *A Natural History of Domesticated Mammals* (Cambridge Univ. Press, 1999).
2. Ducos, P. Proto-élevage et Élevage au Levant Sud au VIIe Millénaire B.C. les Données de la Damascène. *Paléorient* **19,** 153–173 (1993).
3. Garrard, A., Colledge, S. & Martin, L. in *The Origins and Spread of Agriculture and Pastoralism in Eurasia* (ed. Harris, D. R.) 204–226 (University College London Press, 1996).
4. Sherratt, A. in *Patterns of the Past. Studies in Honour of David Clark* (eds Hodder, I., Isaac, G. & Hammond, N.) 261–305 (Cambridge Univ. Press, 1981).
5. Sherratt, A. The secondary exploitation of animals in the Old World. *World Archaeol.* **15,** 90–104 (1983).
6. Dudd, S. N. & Evershed, R. P. Direct demonstration of milk as an element of archaeological economies. *Science* **282,** 1478–1481 (1998).
7. Copley, M. S. *et al.* Direct chemical evidence for widespread dairying in prehistoric Britain. *Proc. Natl Acad. Sci. USA* **100,** 1524–1529 (2003).
8. Copley, M. S. *et al.* Dairying in antiquity: III — Evidence from absorbed lipid residues dating to the British Neolithic. *J. Archaeol. Sci.* **32,** 523–546 (2005).
9. Copley, M. S. *et al.* Processing of milk products in pottery vessels through British prehistory. *Antiquity* **79,** 895–908 (2005).
10. Craig, O. E. *et al.* Did the first farmers of central and eastern Europe produce dairy foods? *Antiquity* **79,** 882–894 (2005).
11. Kohler-Rollefson, I. in *Pastoralism in the Levant: Archaeological Materials in Anthropological Perspective* (eds Bar-Yosef, O. & Khazanov, A.) 11–18 (Monographs in World Archaeology, 1992).
12. Levy, T. E. The emergence of specialized pastoralism in the southern Levant. *World Archaeol.* **15,** 15–36 (1983).
13. Bogucki, P. I. Ceramic sieves of the linear pottery culture and their economic implications. *Oxf. J. Archaeol.* **3,** 15–30 (1984).
14. Payne, S. Kill-off patterns in sheep and goats: The mandibles from Aşvan Kale. *J. Anatol. Stud.* **23,** 281–303 (1973).
15. Legge, A. J. in *Farming Practice in British Prehistory* (ed. Mercer, R.) 169–181 (Edinburgh Univ. Press, 1981).
16. Grigson, C. in *Archaeology of Society in the Holy Land* (ed. Levy, T. E.) 245–268 (Leicester Univ. Press, 1995).
17. Christie, W. W. *Lipid Metabolism in Ruminant Animals* 203–226 (Pergamon, 1981).
18. MacDonald, P., Edwards, R. A. & Greenhalgh, J. F. D. *Animal Nutrition* (Longman, 1988).
19. Evershed, R. P. *et al.* Formation of long-chain ketones in ancient pottery vessels by pyrolysis of acyl lipids. *Tetrahedr. Lett.* **36,** 8875–8878 (1995).
20. Evershed, R. P., Vaughan, S. J., Dudd, S. N. & Coles, J. S. Fuel for thought: Beeswax in lamps and conical cups from Late Minoan Crete. *Antiquity* **71,** 979–985 (1997).
21. Hansel, F. A., Copley, M. S., Madureira, L. A. S. & Evershed, R. P. Thermally produced ω-(*o*-alkylphenyl)alkanoic acids provide evidence for the processing of marine products in archaeological pottery vessels. *Tetrahedr. Lett.* **45,** 2999–3002 (2004).
22. Richards, M. P., Pearson, J. A., Molleson, T. I., Rusell, N. & Martin, L. Stable isotope evidence of diet at Neolithic Catalhöyük, Turkey. *J. Archaeol. Sci.* **30,** 67–76 (2003).
23. Faquhar, G. D., Ehleringer, J. R. & Hubrick, K. T. Carbon isotope discrimination and photosynthesis. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **40,** 503–537 (1989).
24. Boessneck, J. & von den Driesch, A. *Die Tierknockenfunde aus der Neolitischen Siedlung auf dem Fikirtepe bei Kadıköy am Marmaramee* (Institut für Palaeoanatomie, Domestikationsforschung und Geschichte der Tiermedizin, Universität München, 1979).
25. Russell, N. & Martin, L. in *Inhabiting Çatalhöyük: Reports from the 1995–99 Seasons* (ed. Hodder, I.) 33–98 (McDonald Institute Monographs/British Institute of Archaeology, 2005).
26. Mukherjee, A. J., Berstan, R., Copley, M. S., Gibson, A. M. & Evershed, R. P. Compound-specific stable carbon isotope detection of pork consumption applied to the British Late Neolithic. *Antiquity* **81,** 743–754 (2007).
27. Dudd, S. N., Regert, M. & Evershed, R. P. Assessing microbial lipid contributions during laboratory degradations of fats and oils and pure triacylglycerols absorbed in ceramic potsherds. *Org. Geochem.* **29,** 1345–1354 (1998).
28. Aillaud, S. *Field and Laboratory Studies of Diagenetic Reactions Affecting Lipid Residues Absorbed in Unglazed Archaeological Pottery Vessels.* PhD thesis, Univ. Bristol (2000).
29. Charters, S., Evershed, R. P., Goad, L. J., Blinkhorn, P. W. & Denham, V. Quantification and distribution of lipid in archaeological ceramics: Implications for sampling potsherd for organic residue analysis. *Archaeometry* **35,** 211–223 (1993).
30. Mukherjee, A. J., Copley, M. S., Berstan, R. & Evershed, R. P. in *The Zooarchaeology of Fats, Oils, Milk and Dairying* (eds Mulville, J. & Outram, A. K.) 77–92 (Oxbow Books, 2005).

Acknowledgements We thank the Leverhulme Trust for their support (F/00182/T), and the UK Natural Environment Research Council for mass spectrometry facilities.

Author Contributions R.P.E., A.G.S. and S.P. conceived and planned the project. R.P.E. and S.P. wrote the paper. M.S.C., J.C. and D.U.-K. undertook sampling, analytical work and data analysis. All other authors either directed excavations or provided expertise in relation to pottery and/or faunal collections and essential insights into the study region and sites.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to R.P.E. (r.p.evershed@bristol.ac.uk).

LETTERS

Neutralizing antibodies derived from the B cells of 1918 influenza pandemic survivors

Xiaocong Yu[1]*, Tshidi Tsibane[2]*, Patricia A. McGraw[1], Frances S. House[1], Christopher J. Keefer[1], Mark D. Hicar[1], Terrence M. Tumpey[3], Claudia Pappas[2,3], Lucy A. Perrone[3], Osvaldo Martinez[2], James Stevens[3,4], Ian A. Wilson[4], Patricia V. Aguilar[2], Eric L. Altschuler[5], Christopher F. Basler[2] & James E. Crowe Jr[1]

Investigation of the human antibody response to influenza virus infection has been largely limited to serology, with relatively little analysis at the molecular level. The 1918 H1N1 influenza virus pandemic was the most severe of the modern era[1]. Recent work has recovered the gene sequences of this unusual strain[2], so that the 1918 pandemic virus could be reconstituted to display its unique virulence phenotypes[3,4]. However, little is known about adaptive immunity to this virus. We took advantage of the 1918 virus sequencing and the resultant production of recombinant 1918 haemagglutinin (HA) protein antigen to characterize at the clonal level neutralizing antibodies induced by natural exposure of survivors to the 1918 pandemic virus. Here we show that of the 32 individuals tested that were born in or before 1915, each showed seroreactivity with the 1918 virus, nearly 90 years after the pandemic. Seven of the eight donor samples tested had circulating B cells that secreted antibodies that bound the 1918 HA. We isolated B cells from subjects and generated five monoclonal antibodies that showed potent neutralizing activity against 1918 virus from three separate donors. These antibodies also cross-reacted with the genetically similar HA of a 1930 swine H1N1 influenza strain, but did not cross-react with HAs of more contemporary human influenza viruses. The antibody genes had an unusually high degree of somatic mutation. The antibodies bound to the 1918 HA protein with high affinity, had exceptional virus-neutralizing potency and protected mice from lethal infection. Isolation of viruses that escaped inhibition suggested that the antibodies recognize classical antigenic sites on the HA surface. Thus, these studies demonstrate that survivors of the 1918 influenza pandemic possess highly functional, virus-neutralizing antibodies to this uniquely virulent virus, and that humans can sustain circulating B memory cells to viruses for many decades after exposure—well into the tenth decade of life.

Recent studies suggest that the 1918 H1N1 influenza virus was of avian origin[2,5], and is capable of inducing strong systemic cytokine responses that probably contribute to pathogenesis[4,6]. Little is known about naturally occurring adaptive immunity to this virus; however, some elderly survivors are still living. We sought to determine whether survivors showed evidence of acquired immunity to the virus. Expression of the 1918 HA antigen allowed us to identify and characterize protective antibodies induced by natural exposure of humans to the 1918 pandemic virus.

We identified a panel of 32 subjects aged 91–101 years (that is, aged from 2 to 12 in 1918), many of whom recalled a sick family member in the household during the pandemic, which suggested direct exposure to the virus. Of the subjects tested, 100% had serum-neutralizing activity against the 1918 virus (mean titre 1:562), and 94% had serologic reactivity to the 1918 HA (as indicated by haemagglutination inhibition assay (HAI) titres of 1:40 or greater; mean titre 1:396), even though these samples were obtained nearly 90 years after the pandemic. In contrast, subjects born after the pandemic had markedly lower rates of positive serum-neutralizing tests against the 1918 virus (9 out of 10 subjects born 1926–35 had titres $<$1:100, 9 out of 10 subjects born 1936–45 had titres $\leq$1:40, 9 out of 10 subjects born 1946–55 had titres $\leq$1:40). Individual serologic results are shown in Supplementary Table 1.

Peripheral blood mononuclear cells from eight subjects were isolated and B lymphoblastic cell lines were generated by transformation; blood from almost all of the donors tested (7 out of 8) yielded transformed cells secreting antibodies that bound the 1918 HA protein. Supernates from 30 wells of a total of 6,578 wells tested contained 1918 HA-specific antibodies, suggesting a minimal frequency of circulating 1918 HA-specific B cells in the donors of approximately 1 in 4.6×10^6. We collected transformed cells from the wells corresponding to supernates showing the highest levels of specific binding to the 1918 HA (derived from five donors) and fused them to the HMMA2.5 non-secreting myeloma partner[7] using an electrofusion technique[8]. We isolated 17 unique hybridoma cell lines that secreted antibodies reactive with the 1918 HA from cell lines derived from four out of five donors, and then segregated the lines by limiting dilution to yield monoclonal antibody secreting clones. Our screening identified five independent lines with HAI activity against 1918 virus from three separate donors, which we biologically cloned and designated monoclonal antibodies 1I20, 1F1 and 2B12 (donor 6), monoclonal antibody 4D20 (donor 4) and monoclonal antibody 2D1 (donor 23).

Sequence analysis of the antibody genes from the clones demonstrated that the five monoclonal antibodies were distinct and very highly mutated. Genetic features of the antibodies are shown in Table 1. It was of interest that the 1F1, 2B12 and 2D1 clones shared use of the VL1-44*01 gene segment, suggesting a particular fitness for binding of the 1918 virus HA by the CDR1/2 light-chain loops encoded by this VL gene segment. The three clones, however, were clearly independent as they differed in the location of somatic mutations, JL segment (1F1) and in heavy-chain pairing. The numbers of somatic mutations in the variable regions were exceptionally large, almost twice the median number of 18 mutations found in class-switched memory cells in randomly selected human B cells[9]. These data probably suggest recurrent optimization of binding affinity through multiple rounds of somatic hypermutation and selection *in vivo*.

[1]Departments of Pediatrics and of Microbiology and Immunology, Vanderbilt University Medical Center, Nashville, Tennessee 37232, USA. [2]Department of Microbiology, Mount Sinai School of Medicine, New York 10029, USA. [3]Influenza Division, Centers for Disease Control and Prevention, Atlanta, Georgia 30333, USA. [4]Department of Molecular Biology and Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California 92037, USA. [5]Department of Physical Medicine and Rehabilitation, University of Medicine & Dentistry of New Jersey, Newark, New Jersey 07103, USA.
*These authors contributed equally to this work.

Table 1 | Genetic and binding kinetics features of 1918 HA-specific human monoclonal antibodies

	Monoclonal antibody				
	1F1	1I20	2B12	2D1	4D20
Gene segments					
V_H	1-2*02	3-30*02	4-30-4*01	2-70*01	2-26*01
D	5-24*01	3-10*01	3-3*01	1-26*01	4-17*01
J_H	4*02	5*02	4*02	2*01	6*02
V_L	1-44*01	3-15*01	1-44*01	1-44*01	3-21*01
J_L	2*01	1*01	3*02	3*02	1*01
Mutations					
V_H	28	26	32	19†	17
N insertions	2	8	20	15	10
P insertions	1	0	0	0	0
D	1	13	1	2	1
J_H	7	2	1	1	1
V_L	7	9	18	7	14
J_L	3	0	2	0	2
V_L–J_L junction	3	6	6	2	0
Isotype/subclass	IgG1	IgG1	IgG2	IgG1	IgG1
Light chain	λ	κ	λ	λ	λ
Binding kinetics to 1918 HA					
Association rate, K_{on} (ms^{-1})	3.1×10^4	2.9×10^5	1.8×10^4	3.9×10^4	1.7×10^5
Dissociation rate, K_{off} (s^{-1})	1.7×10^{-4}	1.4×10^{-5}	1.2×10^{-4}	1.0×10^{-5}	1.0×10^{-5}
Affinity, K_d (M)	5.4×10^{-9}	4.8×10^{-11}	6.2×10^{-9}	2.5×10^{-10}	1.4×10^{-10}

† Includes a naturally occurring nine-nucleotide insertion starting from the V_H region codon 67.

Purified monoclonal antibodies were assessed by ELISA against a series of representative twentieth century H1N1 viruses including human isolates from 1918, 1943, 1947, 1977 and 1999. We also examined reactivity with influenza A/Swine/Iowa/15/30 (H1N1) virus, as the HA sequence of this virus more closely resembles the 1918 HA sequence than the sequence of any other existing isolate does. The monoclonal antibodies bound to the 1918 HA, with clear cross-reactivity with the 1930 strain, suggesting the remote origin of the antibodies (Fig. 1 and Supplementary Fig. 1). The 1F1 clone also bound to a minimal degree to the 1977 strain, a virus that is almost identical to isolates from the early 1950s[10], and minimally to a 1943 isolate, but not to other post-1930 strains. The antibodies also bound to the corresponding HA expressed on the surface of mammalian cells following transfection of a complementary DNA encoding the 1918 HA as detected by immunofluorescence microscopy (Supplementary Fig. 2). The antibodies did not bind to influenza H3, B or H5 HA proteins in ELISA (data not shown). All five antibodies proved to have very high affinities for recombinant 1918 HA protein when tested by surface plasmon resonance, ranging from 5.4×10^{-9} to 4.8×10^{-11} M (Table 1).

We tested five purified monoclonal antibodies from three separate donors for inhibitory activity in an HAI assay using 1918 virus-like particles (VLP) or a panel of H1N1 viruses. The antibodies showed specific binding for 'old' viruses, including the 1918 virus or viruses that were genetically similar to the 1918 virus (Table 2 and Supplementary Table 2). Specifically, all five antibodies reacted with 1918 HA by ELISA, by HAI with VLP of the 1918 strain or with the highly genetically similar influenza A/Swine/Iowa/15/30 (H1N1) virus (Sw/30), and by neutralizing assay with reconstituted 1918 virus. Comparable HAI activities were obtained for both 1918 and Sw/30 VLP in these assays (Table 2). The five clones had specific neutralizing activities ranging from 0.32 to 0.97 $\mu g\,ml^{-1}$ and HAI activities of 0.18 to 0.47 $\mu g\,ml^{-1}$ against the 1918 virus. In contrast, the same antibodies failed to interact with, or inhibit, human H1N1 viruses isolated in 1943, 1947 or 1999 (Supplementary Table 2). The 1F1 antibody bound and neutralized the 1977 virus, albeit to a lesser degree than either the 1918 or the Sw/30 viruses (showing a specific HAI activity of 0.16 $\mu g\,ml^{-1}$ against the 1977 virus), and to a minimal degree the 1943 virus (Supplementary Table 2). Again, only 1F1 displayed neutralizing activity against the 1943 or 1977 viruses, with specific activities of 1.8 and 0.88 $\mu g\,ml^{-1}$, respectively.

We selected antibody escape mutants for three monoclonal antibodies using the Sw/30 virus. Nucleotide sequence analysis of the HA genes from these viruses demonstrated that they had acquired mutations in classical antigenic regions of HA. The established antigenic map of influenza A (H1N1) viruses specifies an amino acid numbering scheme and indicates five immunodominant antigenic sites, designated Sa, Sb, Ca_1, Ca_2 and Cb[11,12]. The 2B12 mutants possessed mutations at residue 166 (K166Q, K166E or K166P), which lies in the Sa antigenic site (Supplementary Fig. 3), using the previous numbering scheme[11]. Escape mutants for 1F1 and 1I20 possessed an identical mutation, P186H, at a residue adjacent to the previously

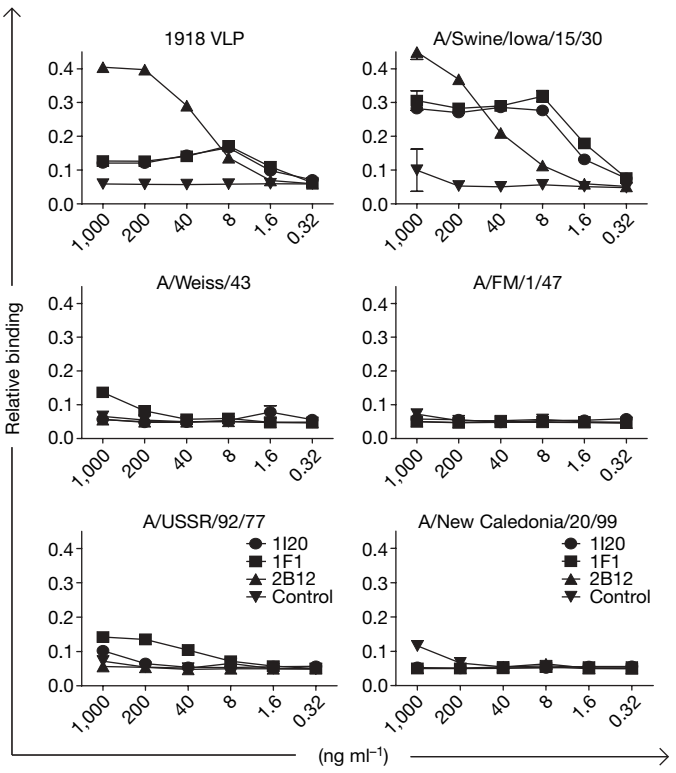


Figure 1 | Binding of human monoclonal antibodies to representative twentieth century H1N1 viruses. Equivalent HA units of 1918 VLP, A/Swine/Iowa/15/30, A/Weiss/43, A/FM/1/47, A/USSR/92/77 and A/New Caledonia/20/99 influenza viruses were absorbed onto ELISA plates. An ELISA was performed using serial 1:5 dilutions of monoclonal antibodies 1F1, 1I20, 2B12 or an H5-specific control human monoclonal antibody. Relative binding (*y* axis) indicates optical density in ELISA binding assay to absorbed VLP or virus.

Table 2 | Specific HAI activity of human antibodies against influenza viruses or VLP

Antigen	Monoclonal antibody				
	1F1	1I20	2B12	2D1	4D20
Sw/30 viruses					
Wild type	0.04	0.08	1.25	0.08	0.04
P186H (site Sb)	–	–	1.25	ND	ND
K166E (site Sa)	0.04	0.04	–	ND	ND
K166Q (site Sa)	0.04	0.04	–	ND	ND
1918 VLP					
Wild type	0.08	0.16	0.31	0.01	0.04
P186H (site Sb)	–	0.63	0.31	0.01	0.04
K166E (site Sa)	<0.04	0.08	–	–	0.63
K166Q (site Sa)	<0.04	0.04	–	2.5	0.31

Specific HAI activity ($\mu g\,ml^{-1}$) of monoclonal antibodies was calculated as the lowest concentration of antibody that displayed haemagglutinating activity. – , activity was not detected at any concentration tested, up to 2.5 $\mu g\,ml^{-1}$; ND, not determined.

defined Sb antigenic site (Supplementary Fig. 3), which encompasses the α-helix of the receptor binding site in the 1918 HA structure[11–13]. As expected, incorporation of the Sb antigenic site mutation into VLP reduced (1I20) or eliminated (1F1) activity of the corresponding antibodies in an HAI assay without affecting binding of 2B12 (Table 2). The 2D1 and 4D20 antibodies also had reduced HAI activity against the site Sb mutant viruses and VLP, suggesting that they also bound to this site. In contrast, incorporation of the Sa mutation into VLP abolished activity of 2B12 in the HAI assay, but did not affect the activity of the other four antibodies (Table 2). The 1918 and Sw/30 virus HAs differ by only a single amino acid in each of the Sa and Sb sites (Supplementary Fig. 3), explaining why these antibodies cross-neutralize both viruses. In contrast, the 1943 isolate contains seven changes in the Sa site and seven changes in the Sb site, relative to the 1918 HA sequence, explaining the partial or complete loss of neutralizing ability of the antibodies for these later isolates. Notably, 9 of the 12 residues at the Sb site differed between the 1918 and the 1977 virus. The ability of 1F1 to cross-neutralize the 1977 virus despite its divergence from the 1918 virus is worthy of further investigation, because these data raise the possibility that 1F1 recognizes a new, more broadly neutralizing epitope. Alternatively, the data may represent epitope recycling such that, despite significant sequence divergence, the 1F1 epitope reappeared. In either case, further characterization of this antibody may suggest strategies to elicit enhanced cross-protective immunity to influenza A viruses of a particular HA subtype.

We tested the five antibodies for therapeutic efficacy in an established mouse model of infection. Mice were inoculated by the intranasal route with the previously reconstructed 1918 virus, and morbidity (measured by weight loss), mortality and virus replication were assessed as previously described[3]. Each of the five 1918-specific monoclonal antibodies tested showed therapeutic efficacy when administered 1 day after infection, preventing death of animals (Table 2). Mice treated with a control (H5 HA-specific) human monoclonal antibody or human IgG did not survive. Reduced weight loss and lower levels of virus replication in the lungs of anti-1918-antibody-treated mice on day 4 after infection also demonstrated a significant protective effect that correlated well with survival data. At lower doses, the antibodies caused a statistically significant delay to death, relative to the controls.

These studies suggest that B cells responding to viral infections, or their progeny, survive for the life of the host, even nine or more decades after exposure. It is well established that a subset of plasma cells is long-lived[14], and these cells contribute to durable humoral immune responses[15], such as that observed after childhood smallpox vaccination[16–19]. Memory B cell pools also can be long-lived,

Table 3 | Therapeutic efficacy of 1918 HA-specific monoclonal antibodies against 1918 virus in mice

Antibody	Experiment	Dose, per mouse (μg)*	Weight loss (%)†	Virus in lung‡	Survival (no. protected/ total no.)§
2B12	1	200	2.3	3.8 ± 1.4	5/5
		20	11	ND	5/5
		2	14	ND	0/5
1F1	1	200	1.0	3.9 ± 1.4	5/5
		20	12	ND	4/5
		2	10	ND	0/5
1I20	1	20	4.2	5.4 ± 0.2	5/5
		2	14	ND	0/5
		0.2	16	ND	0/5
2D1	2	200	3.0	3.3 ± 0.5	5/5
		20	10	5.1 ± 0.2	5/5
		2	14	5.9 ± 0.5	0/5
4D20	2	200	7.6	3.5 ± 0.6	5/5
		20	15	5.4 ± 0.5	2/5
		2	17	5.8 ± 0.4	0/5
Control H5 HA antibody	1	200	13	6.4 ± 0.3	0/5
		20	15	ND	0/5
		2	18	ND	0/5
Control human IgG	1	200	15	6.3 ± 0.2	0/5
	2		18	6.4 ± 0.3	0/5
	1	20	11	ND	0/5
	2		16	ND	0/5
	1	2	14	ND	0/5
	2		16	ND	0/5

The statistical significance of viral titre and morbidity data was determined by analysis of variance (ANOVA). The statistical significance of mortality data was determined using the null model likelihood ratio test and the Mann–Whitney *U*-test.

* Groups of mice were infected intranasally with five LD_{50} of 1918 virus and then treated 24 h later with graded doses of 1918-specific monoclonal antibodies or control antibodies.

† Maximum per cent weight loss (mean of five mice per group); highest antibody dose of each experimental group, $P \leq 0.05$ (ANOVA) versus control antibody groups.

‡ Average lung titres of three mice on day 4 after inoculation, expressed as (log_{10}) 50% egg infectious dose (EID_{50}) per ml ± s.d. All 20 and 200 μg dose experimental groups, $P \leq 0.024$ (ANOVA) versus control antibody groups.

§ High antibody dose of each experimental group and 20 μg of 1F1 or 2D11, $P \leq 0.0031$ (Mann–Whitney *U*-test) versus control antibody groups. ND, not determined.

sustained in part by antigen-independent polyclonal stimuli[20]. It is difficult to be absolutely certain that the monoclonal antibodies isolated here were first stimulated by exposure during the 1918 pandemic. However, the clinical history of the subjects and the high functional specificity of the monoclonal antibodies for the 1918 strain strongly suggest that recent exposures do not account for this immunity. Probably, boosting by antigenically related viruses in the early decades of the twentieth century may have contributed to the ability of these subjects to sustain these B cells. The variable genes of five independent human neutralizing antibodies had a very high frequency of somatic mutations, associated with strong binding constants and high potency. The *in vivo* efficacy of treatment with these antibodies shows that the development of functional adaptive immunity to the pandemic virus did occur in survivors of the 1918 pandemic.

It has long been known that infusion of neutralizing antibodies can protect mice from lethal influenza virus infection, and transfusion of convalescent blood products to 1918 influenza victims may have had a beneficial effect[21]. Thus, the monoclonal antibodies described here could serve as potential therapeutics for a re-emergent 1918-like virus. The techniques described here suggest that it may be possible to recover human antibodies that display a wide array of specificities corresponding to the viruses and other pathogens that have infected an individual during their lifetime.

METHODS SUMMARY

Recombinant 1918 virus HA (A/South Carolina/1/1918) was produced as described[11]. Peripheral blood mononuclear cells were obtained from volunteers born in 1915 or earlier. Hybridomas were generated from EBV-transformed B cell lines by electrofusion to the HMMA2.5 cell line[7,8]. When hybridoma lines formed colonies in the presence of selecting drugs, lines were cloned by limiting dilution. Secreted monoclonal antibodies were concentrated and purified by fast protein liquid chromatography (FPLC). The isotype and subclass of secreted antibodies were determined by ELISA. Nucleotide sequences of variable gene segments were determined by automated sequence analysis of cloned cDNA[22]. The identity of the gene segments and mutations from the germline sequences were determined by alignment using the ImMunoGeneTics database[23]. Viruses were propagated in 10-day-old embryonated chicken eggs. Influenza A/South Carolina/1/18 virus was prepared as previously described[3]. Expression plasmids encoding the 1918 HA and neuraminidase (NA) proteins were described previously[24,25]. Binding of antibodies was determined using 1918 VLP or influenza A viruses as the coating antigen in ELISA. VLP were produced by co-transfection of 293T cells with expression plasmids for the 1918 HA and 1918 NA, consistent with a recent report[26]. HAI assays of sera or antibodies were performed according to standard protocols using chicken red blood cells[27]. For microneutralization assay, ten 50% tissue-culture infective dose units ($TCID_{50}$) of virus was preincubated with dilutions of sera or monoclonal antibody and then used to infect Madin–Darby canine kidney (MDCK) cells in 96-well plates, as described[28,29]. The kinetic interaction of monoclonal antibodies with the 1918 HA protein was determined by surface plasmon resonance. Antibody escape mutants were isolated by treatment of Sw/30 virus with excess antibody as described[12,30]. Mice were inoculated intranasally with five LD_{50} (lethal dose to 50% of animals) of the 1918 virus. At 24 h after inoculation, we administered 1918-specific monoclonal antibody or control antibodies to each mouse. Mice were observed for weight loss or death. Subsets of animals were killed for virus titre.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 7 April; accepted 3 July 2008.
Published online 17 August 2008.

1. Johnson, N. P. & Mueller, J. Updating the accounts: global mortality of the 1918–1920 "Spanish" influenza pandemic. *Bull. Hist. Med.* **76,** 105–115 (2002).
2. Taubenberger, J. K. *et al.* Characterization of the 1918 influenza virus polymerase genes. *Nature* **437,** 889–893 (2005).
3. Tumpey, T. M. *et al.* Characterization of the reconstructed 1918 Spanish influenza pandemic virus. *Science* **310,** 77–80 (2005).
4. Kobasa, D. *et al.* Aberrant innate immune response in lethal infection of macaques with the 1918 influenza virus. *Nature* **445,** 319–323 (2007).
5. Taubenberger, J. K. The origin and virulence of the 1918 "Spanish" influenza virus. *Proc. Am. Phil. Soc.* **150,** 86–112 (2006).
6. Kash, J. C. *et al.* Genomic analysis of increased host immune and cell death responses induced by 1918 influenza virus. *Nature* **443,** 578–581 (2006).
7. Posner, M. R., Elboim, H. & Santos, D. The construction and use of a human-mouse myeloma analogue suitable for the routine production of hybridomas secreting human monoclonal antibodies. *Hybridoma* **6,** 611–625 (1987).
8. Yu, X., McGraw, P. A., House, F. S. & Crowe, J. E. Jr. An optimized electrofusion-based protocol for generating virus-specific human monoclonal antibodies. *J. Immunol. Methods* **336,** 142–151 (2008).
9. Tian, C. *et al.* Evidence for preferential Ig gene usage and differential TdT and exonuclease activities in human naive and memory B cells. *Mol. Immunol.* **44,** 2173–2183 (2007).
10. Nakajima, K., Desselberger, U. & Palese, P. Recent human influenza A (H1N1) viruses are closely related genetically to strains isolated in 1950. *Nature* **274,** 334–339 (1978).
11. Stevens, J. *et al.* Structure of the uncleaved human H1 hemagglutinin from the extinct 1918 influenza virus. *Science* **303,** 1866–1870 (2004).
12. Caton, A. J., Brownlee, G. G., Yewdell, J. W. & Gerhard, W. The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). *Cell* **31,** 417–427 (1982).
13. Brownlee, G. G. & Fodor, E. The predicted antigenicity of the haemagglutinin of the 1918 Spanish influenza pandemic suggests an avian origin. *Phil. Trans. R. Soc. Lond. B* **356,** 1871–1876 (2001).
14. Manz, R. A., Thiel, A. & Radbruch, A. Lifetime of plasma cells in the bone marrow. *Nature* **388,** 133–134 (1997).
15. Slifka, M. K., Antia, R., Whitmire, J. K. & Ahmed, R. Humoral immunity due to long-lived plasma cells. *Immunity* **8,** 363–372 (1998).
16. Amanna, I. J., Slifka, M. K. & Crotty, S. Immunity and immunological memory following smallpox vaccination. *Immunol. Rev.* **211,** 320–337 (2006).
17. Crotty, S. *et al.* Cutting edge: long-term B cell memory in humans after smallpox vaccination. *J. Immunol.* **171,** 4969–4973 (2003).
18. Hammarlund, E. *et al.* Duration of antiviral immunity after smallpox vaccination. *Nature Med.* **9,** 1131–1137 (2003).
19. Amanna, I. J., Carlson, N. E. & Slifka, M. K. Duration of humoral immunity to common viral and vaccine antigens. *N. Engl. J. Med.* **357,** 1903–1915 (2007).
20. Bernasconi, N. L., Traggiai, E. & Lanzavecchia, A. Maintenance of serological memory by polyclonal activation of human memory B cells. *Science* **298,** 2199–2202 (2002).
21. Luke, T. C., Kilbane, E. M., Jackson, J. L. & Hoffman, S. L. Meta-analysis: convalescent blood products for Spanish influenza pneumonia: a future H5N1 treatment? *Ann. Intern. Med.* **145,** 599–609 (2006).
22. Weitkamp, J. H. *et al.* Generation of recombinant human monoclonal antibodies to rotavirus from single antigen-specific B cells selected with fluorescent virus-like particles. *J. Immunol. Methods* **275,** 223–237 (2003).
23. Ruiz, M. *et al.* IMGT, the international ImMunoGeneTics database. *Nucleic Acids Res.* **28,** 219–221 (2000).
24. Tumpey, T. M. *et al.* Existing antivirals are effective against influenza viruses with genes from the 1918 pandemic virus. *Proc. Natl Acad. Sci. USA* **99,** 13849–13854 (2002).
25. Glaser, L. *et al.* A single amino acid substitution in 1918 influenza virus hemagglutinin changes receptor binding specificity. *J. Virol.* **79,** 11533–11536 (2005).
26. Chen, B. J., Leser, G. P., Morita, E. & Lamb, R. A. Influenza virus hemagglutinin and neuraminidase, but not the matrix protein, are required for assembly and budding of plasmid-derived virus-like particles. *J. Virol.* **81,** 7111–7123 (2007).
27. World Health Organization Collaborating Centers for Reference and Research on Influenza. in *Concepts and Procedures for Laboratory-Based Influenza Surveillance* (eds Kendal, A. P., Skehel, J. J. & Pereira, M. S.) B17–B35 (Centers for Disease Control and Prevention, 1982).
28. Mozdzanowska, K. *et al.* A pulmonary influenza virus infection in SCID mice can be cured by treatment with hemagglutinin-specific antibodies that display very low virus-neutralizing activity *in vitro*. *J. Virol.* **71,** 4347–4355 (1997).
29. Reed, L. J. & Muench, H. A simple method of estimating fifty percent endpoints. *Am. J. Hyg.* **27,** 493–497 (1938).
30. Yewdell, J. W., Webster, R. G. & Gerhard, W. U. Antigenic variation in three distinct determinants of an influenza type A haemagglutinin molecule. *Nature* **279,** 246–248 (1979).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank L. Schoenherr, K. Sharma, L. Adams (supported by the University of Medicine and Dentistry of New Jersey (UMDNJ) Institute for the Elimination of Health Disparities), C. Dokes, C. Gaines, B. Butter and S. Rivera for assistance with subjects, and S. Yoder and B. Briney for sample preparation. This work was supported by grants from the National Institutes of Health to J.E.C. (U54 AI057157, Southeast Regional Center of Excellence for Emerging Infections and Biodefense, and U19 AI057229), and C.F.B. (U54 AI57158, Northeast Biodefense Center, U19 AI62623, Center for Investigating Viral Immunity and Antagonism, and P01 AI058113). I.A.W. and J.S. were supported in part by the National Institutes of Health (CA55896 and AI42266). P.V.A. was supported by a fellowship awarded by the Northeast Biodefense Center (AI057158). We thank P. Palese and A. Garcia-Sastre for advice and for providing H1N1 viruses, and M. Posner and L. Cavacini for the HMMA2.5 cell line.

Author Contributions X.Y., P.A.M., M.D.H. and F.S.H. made and cloned the monoclonal antibodies, sequenced antibody genes, and performed immunofluorescence experiments. T.T. characterized the interaction of the antibodies with viruses and VLPs and selected for and characterized the escape mutants. C.J.K. performed biosensor studies. T.M.T., C.P. and L.A.P. designed and performed *in vivo* studies. O.M. sequenced the HA genes of the H1N1 viruses used in this study and performed ELISA assays with these viruses. P.V.A. assisted with HAI and neutralization assays and with cloning of recombinant HA molecules. J.S. and I.A.W. provided recombinant HA. E.L.A. led the clinical recruitment and, together with C.F.B. and J.E.C., conceived of the experimental plan. C.F.B. and J.E.C. wrote the manuscript. All authors discussed the results and commented on the manuscript.

Author Information Antibody nucleotide sequences have been deposited in GenBank under accession numbers EU169674 to EU169679, and EU825947 to EU825950. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.E.C. (James.Crowe@vanderbilt.edu), C.F.B. (Chris.Basler@mssm.edu), or E.L.A. (altschel@umdnj.edu).

METHODS

Subjects. Volunteers born in 1915 or earlier were recruited at the UMDNJ site. We obtained 50 ml of peripheral venous blood from each subject after informed written consent.

Antigen. Recombinant A/South Carolina/1/1918 virus HA was generated in a baculovirus expression system and purified, as described[11].

Production of monoclonal antibodies. Peripheral blood monocytes were obtained from eight donors by density gradient centrifugation of whole heparinized blood. B cells were transformed in 384-well plates with approximately 1,000 B cells per well by *in vitro* culture in medium with CpG ODN 10103 (Coley) and Epstein–Barr virus (EBV) (supernate from cell line B95.8). Supernates from the resulting cell lines were tested for reactivity with the 1918 HA by ELISA. Hybridomas were generated using cells from wells containing reactive lines by fusion to the HMMA2.5 non-secreting myeloma cell line[7] using electrofusion followed by HAT (hypoxanthine, aminopterin and thymidine) and ouabain drug selection, as described[8]. When hybridomas formed colonies, lines were screened using an anti-1918 HA ELISA; positive lines were cloned by limiting dilution. Supernates of high-density cultures were produced in CELLine devices (BD Biosciences). Purified and concentrated preparations of each of the antibodies were prepared by FPLC using protein G conjugated resin on an AKTA instrument (GE Healthcare) followed by concentration and buffer exchange using ultra centrifugal filter devices (Millipore).

Characterization of monoclonal antibodies. The isotype and subclass of secreted antibodies were determined by ELISA. Nucleotide sequences of variable gene segments were determined from cloned cDNA generated by PCR with reverse transcription (RT–PCR) amplification of cellular messenger RNA using variable gene-specific primers designed to amplify antibody genes from all gene families[22]. Identity of the gene segments and mutations from the germline sequences were determined by alignment using the ImMunoGeneTics database (http://imgt.cines.fr)[23].

Viruses and plasmids. The following viruses were propagated in 10-day-old embryonated chicken eggs: influenza A/Weiss/43, A/FM/1/47, A/USSR/92/77, A/New Caledonia/20/99, A/Swine/Iowa/15/30. The sequences of the HA genes of these viruses used were confirmed by RT–PCR and cDNA sequence analysis. Influenza A/South Carolina/1/18 virus was prepared as described[3]. Expression plasmids encoding the 1918 HA and NA proteins were described previously[24,25]. The 1918 virus was handled under biosafety level 3 enhanced (BSL3) containment in accordance with guidelines of the National Institutes of Health (NIH) and the Centers for Disease Control and Prevention (CDC) (available at http://www.cdc.gov/flu/h2n2bsl3.htm) and in accordance with requirements of the US Department of Agriculture (USDA)-CDC select agent program.

ELISA. Equivalent HA units of 1918 VLP or of influenza A viruses were diluted in coating buffer (51-2713KC; BD Biosciences) and adsorbed overnight onto ELISA plates (Nunc). Plates were washed with PBS containing 0.05% Tween-20 and blocked using blocking solution (555213; BD Biosciences) or PBS containing 5% FCS at room temperature for 1 h. Diluted monoclonal antibodies were added to the plates and incubated at 4 °C for 16 h. Plates were washed and incubated with HRP-conjugated goat anti-human IgG (H10507; Caltag) for 1 h. Washed plates were developed using TMB substrate, and the reaction was stopped using 1 M sulphuric acid. Optical density was read at 450 nm using an ELISA plate reader.

Production of VLP. VLP were produced by co-transfection of 10^6 293T cells with 1 μg each of expression plasmids for the 1918 HA and NA. Two days after transfection, supernatants were collected and assayed for HA activity. Our ability to produce these haemagglutinating particles in the absence of other viral proteins is consistent with a recent report[26].

HAI assays of sera or monoclonal antibodies were performed according to standard protocols[27]. In brief, sera were initially diluted 1:10 in receptor-destroying enzyme from *Vibrio cholerae* (Denka Seiken). Serial dilutions of sera or monoclonal antibodies were pre-incubated with 8 HA units of virus per well. Chicken red blood cells were added to a final concentration of 0.5% and the plate was incubated on ice for 30–60 min.

Microneutralization assay. Ten $TCID_{50}$ units of virus were preincubated with dilutions of sera or monoclonal antibody and used to infect MDCK cells in 96-well plates with six replicates, as described[28]. Neutralizing concentrations were defined as the reciprocal of the highest dilution of serum in which 50% of wells were infected, as calculated by the method of ref. 29. Specific neutralizing activity of antibodies was calculated as the lowest concentration of antibody that displayed activity.

Biosensor studies. The kinetic interaction of antibodies with recombinant 1918 HA protein was determined by surface plasmon resonance using a Biacore 2000 instrument. Purified 1918 HA protein was diluted to 30 μg ml^{-1} in 10 mM sodium acetate, pH 4.5, and covalently immobilized at 5 μl min^{-1} by amine coupling to the dextran matrix of a CM5 sensor chip (Biacore AB) with a target density of 1,200 response units. Unreacted active ester groups were blocked with 1 M ethanolamine. All five purified 1918 monoclonal antibodies and a human H5-specific influenza antibody (negative control) at concentrations ranging from 5 to 500 nM in HBS/Tween-20 buffer (Biacore AB) were injected over the immobilized 1918 HA protein or reference cell surface. Association rates (K_{on}), dissociation rates (K_{off}), and affinities or equilibrium dissociation constants (K_d) were calculated by aligning the binding curves globally to fit a 2:1 Langmuir binding model using BIAevaluation 4.1 software. The goodness of each fit was based on the agreement between experimental data and the calculated fits, in which the X^2 values were below 1.0.

Selection and characterization of antibody escape mutants. Mutants were isolated as described[12,30]. In brief, escape mutant viruses were selected by treatment of Sw/30 virus with excess antibody, followed by recovery of neutralization-resistant viruses in eggs. RNA was extracted from virus-infected allantoic fluid, then cDNA was generated by RT–PCR, cloned, sequenced and aligned to previously determined wild-type virus HA gene sequences.

Animal studies. Female BALB/c (8-week-old) mice were inoculated intranasally with 5 LD_{50} in a 50 μl volume of the virulent reconstituted 1918 virus. At 24 h after inoculation, we administered 200, 20, 2 or 0.2 μg (approximately 10, 1, 0.1 or 0.01 mg kg^{-1}) of 1918-specific monoclonal antibody or a similarly prepared human monoclonal antibody to H5 influenza HA (clone IE5), or an equal volume of human IgG, to each mouse, in groups of 11 (highest dose) or 5 (lower doses) mice. Mice were observed for 16 days for weight loss or death. Subsets of three animals treated with the highest dose were killed on day 2 and 4 after infection, and whole lungs were homogenized in 1 ml of sterile PBS. Virus titre in lung tissue homogenates was determined by plaque titration in MDCK cell monolayer cultures.

LETTERS

Functional auditory hair cells produced in the mammalian cochlea by *in utero* gene transfer

Samuel P. Gubbels[1]*†, David W. Woessner[1]*†, John C. Mitchell[2], Anthony J. Ricci[3] & John V. Brigande[1]

Sensory hair cells in the mammalian cochlea convert mechanical stimuli into electrical impulses that subserve audition[1,2]. Loss of hair cells and their innervating neurons is the most frequent cause of hearing impairment[3]. Atonal homologue 1 (encoded by *Atoh1*, also known as *Math1*) is a basic helix–loop–helix transcription factor required for hair-cell development[4–6], and its misexpression *in vitro*[7,8] and *in vivo*[9,10] generates hair-cell-like cells. *Atoh1*-based gene therapy to ameliorate auditory[10] and vestibular[11] dysfunction has been proposed. However, the biophysical properties of putative hair cells induced by *Atoh1* misexpression have not been characterized. Here we show that *in utero* gene transfer of *Atoh1* produces functional supernumerary hair cells in the mouse cochlea. The induced hair cells display stereociliary bundles, attract neuronal processes and express the ribbon synapse marker carboxy-terminal binding protein 2 (refs 12,13). Moreover, the hair cells are capable of mechanoelectrical transduction[1,2] and show basolateral conductances with age-appropriate specializations. Our results demonstrate that manipulation of cell fate by transcription factor misexpression produces functional sensory cells in the postnatal mammalian cochlea. We expect that our *in utero* gene transfer paradigm will enable the design and validation of gene therapies to ameliorate hearing loss in mouse models of human deafness[14,15].

We devised an *in utero* gene transfer method to conduct gain-of-function studies in the developing mouse inner ear[16] (Fig. 1 and Supplementary Figs 1 and 2). A plasmid consisting of the human elongation factor 1-α gene (*EEF1A1*; also known as *EF1A*) promoter[17] driving the expression of a destabilized form (2 h half-life) of green fluorescent protein (ZsGreen)[18,19] was microinjected through the uterus into the fluid-filled cavity of the embryonic day 11.5 (E11.5) mouse otic vesicle (Fig. 1a and Supplementary Video 1). The plasmid-filled left otocyst was centred in the field of a paddle-style circular electrode, and a directional square-wave pulse train was delivered to electroporate ventral progenitor cells that give rise to the organ of Corti (Fig. 1b and Supplementary Video 2). ZsGreen was detected 24 h after electroporation in a teardrop-shaped pattern consistent with the morphology of the otocyst (four of six otocysts; Fig. 1c and Supplementary Fig. 2). Histological analysis revealed ZsGreen-positive progenitor cells in their stereotyped pseudostratified arrangement within the otic epithelium (Fig. 1d)[20]. These data indicate that *in vivo* electroporation transfects otic epithelial progenitor cells in the ventromedial otocyst and produces robust transgene expression within 24 h.

To determine which differentiated cell types arise from transfected otic progenitors, we injected and electroporated EF1α-enhanced green fluorescent protein (GFP) at E11.5 and analysed E18.5 cochlear whole mounts immunostained for the hair-cell marker myosin 7a (Myo7a)[21,22]. The gross morphology of the inner ear at E18.5 was unaffected, and electroporated embryos carried to term had normal auditory function one month after birth (Supplementary Table 1). Robust GFP expression was present in inner and outer hair cells and supporting cells in the organ of Corti at E18.5 (Fig. 2a, b, d, f and

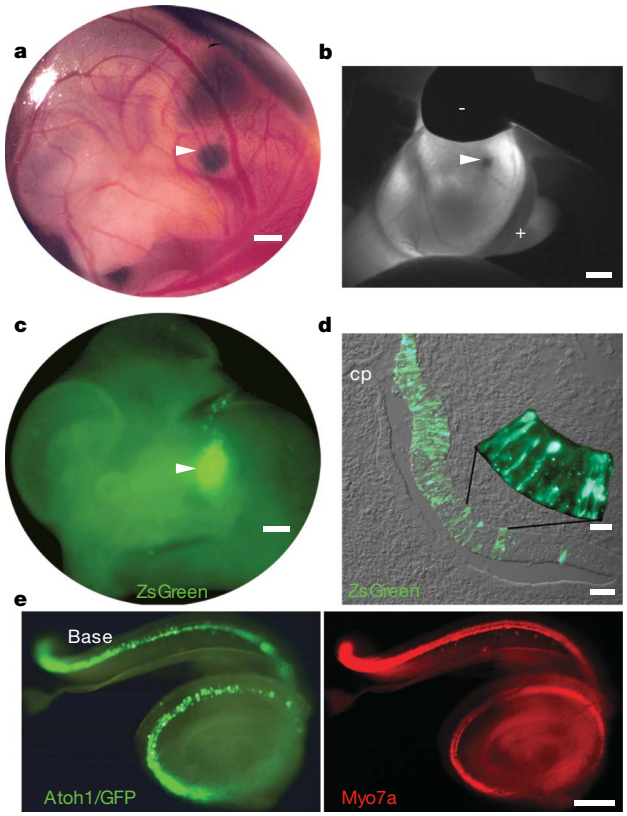


Figure 1 | *In utero* gene transfer to the developing mouse inner ear. **a**, Expression plasmid was microinjected into the E11.5 otic vesicle. **b**, The vesicle was centred between the cathode (−) and anode (+) and electroporated. **c**, Destabilized GFP (ZsGreen) was expressed in the otic territory 24 h after electroporation. **d**, E12.5 progenitors in the medial and ventral otic epithelium expressed ZsGreen robustly. **e**, E18.5 *Atoh1*/GFP-transfected cochlea (left) immunostained for Myo7a (right). Arrowheads indicate left otocyst. cp, lateral canal plate; scale bars, 200 μm (**a**); 500 μm (**b**, **c**); 50 μm (**d**); 10 μm (**d**, inset); 100 μm (**e**).

[1]Department of Otolaryngology, Oregon Hearing Research Center, and [2]Department of Restorative Dentistry, Division of Biomaterials and Biomechanics, School of Dentistry, Oregon Health & Science University, 3181 SW Sam Jackson Park Road, Portland, Oregon 97239, USA. [3]Department of Otolaryngology-Head and Neck Surgery, Stanford University School of Medicine, 801 Welch Road, Stanford, California 94305, USA. †Present addresses: Department of Surgery, Division of Otolaryngology, University of Wisconsin – Madison, K4/719 CSC, 600 Highland Avenue, Madison, Wisconsin 53792, USA (S.P.G.); Department of Pharmacology and Toxicology, University of Utah, College of Pharmacy, 30 South 2000 East, Room 201, Salt Lake City, Utah 84112, USA (D.W.W.).
*These authors contributed equally to this work.

Supplementary Fig. 3). The distribution of transfected cells in the cochlea was constrained to the organ of Corti proper. Differential regulation of GFP expression by the *EEF1A1* promoter may account for this restricted pattern. These data indicate that *in vivo* electroporation transfects otic progenitors that give rise to all of the constituent cell types within the organ of Corti but it does not adversely affect gross embryonic development or the postnatal acquisition of hearing.

To induce otic epithelial progenitor cells to adopt a hair-cell fate, we misexpressed *Atoh1*, a basic helix–loop–helix transcription factor required for hair-cell formation[4]. A plasmid generating a bicistronic message encoding Atoh1 and enhanced GFP was injected and electroporated in the E11.5 otocyst. The gross distribution of Atoh1/GFP expression in the E18.5 cochlea seemed to follow that defined by Myo7a (Fig. 1e), suggesting that progenitors giving rise to the organ of Corti were transfected with *Atoh1*. Atoh1/GFP$^+$ cells co-expressing Myo7a were present in the base, midbase and apex of transfected cochleae (Fig. 2c, e, g). The stereotyped pattern of one inner and three outer rows of hair cells was altered by the overabundance of Atoh1/GFP$^+$/Myo7a$^+$ cells that we refer to as supernumerary cells (Fig. 2c, e, g). The apical surfaces of the supernumerary cells had phalloidin-positive epithelial protrusions at E18.5 that resembled immature stereociliary bundles (Supplementary Fig. 4a–d), which persisted for one month after birth (Fig. 2h–k). Ectopic Myo7a$^+$ cells displaced

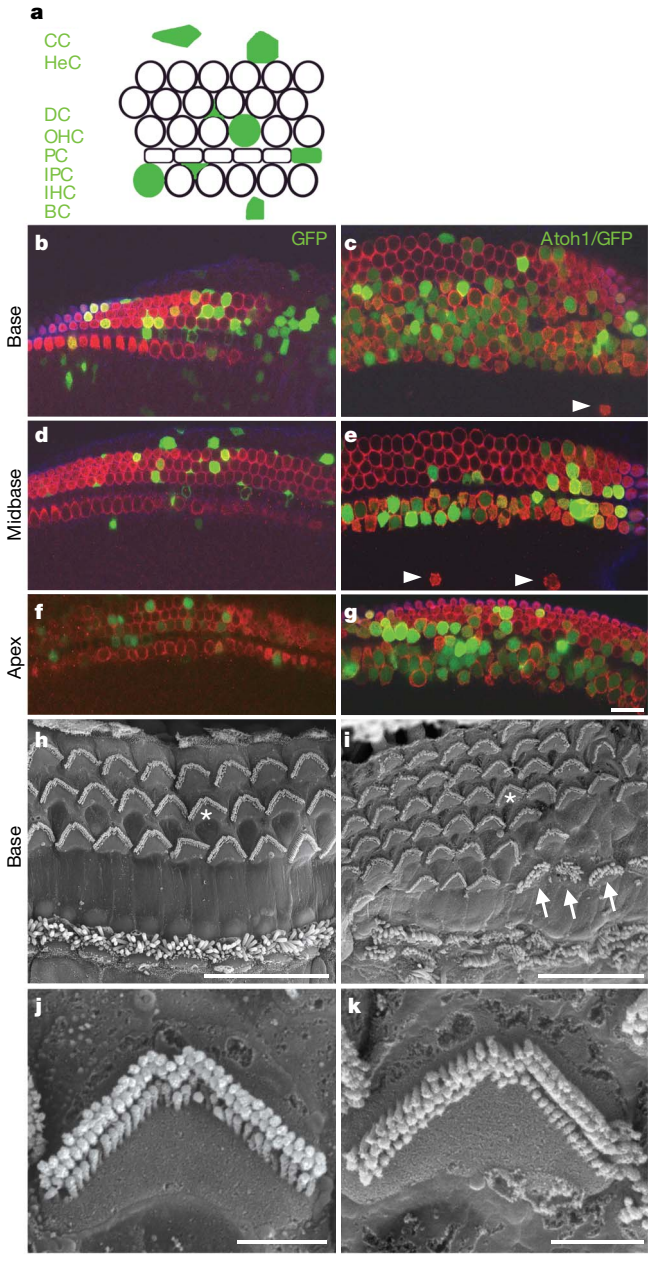


Figure 2 | *Atoh1* misexpression generates supernumerary Myo7a$^+$ cells bearing stereociliary bundles. **a**, Schematic diagram of cell types in the organ of Corti transfected with *Atoh1*/enhanced GFP (filled green): CC, Claudius's cells; HeC, Hensen's cells; DC, Deiters's cells; PC, pillar cell; IPC, inner phalangial cell; BC, border cell. **b**–**g**, Laser confocal micrographs of E18.5 GFP-transfected (**b**, **d**, **f**) and *Atoh1*/GFP-transfected (**c**, **e**, **g**) organs of Corti immunostained for Myo7a (red). All of the Atoh1/GFP$^+$ cells in **c**, **e** and **g** are Myo7a$^+$. **h**–**j**, scanning electron micrographs of postnatal day 35, untransfected (**h**) and *Atoh1*/GFP-transfected (**i**) organs of Corti. Asterisks in **h** and **i** indicate stereociliary bundles imaged at higher magnification in **j** and **k**, respectively. The arrows indicate three cells with atypical bundles. Scale bars, 20 μm (**b**– **i**); 2 μm (**j**, **k**).

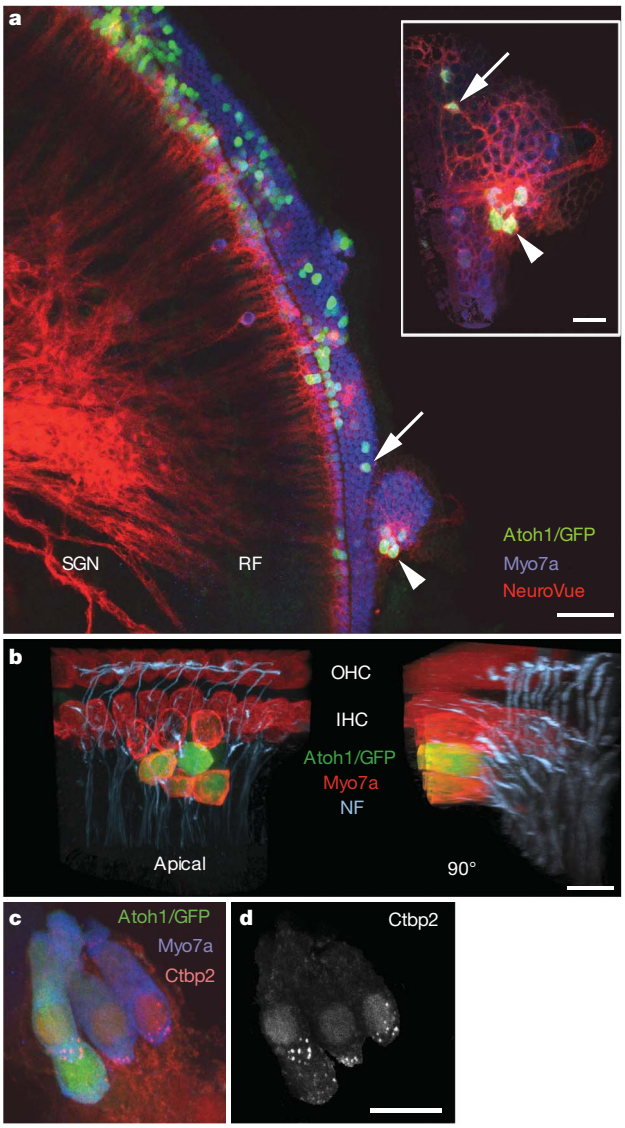


Figure 3 | Atoh1/GFP$^+$ cells show morphological and molecular correlates of innervation and synaptogenesis. **a**, NeuroVue-Red placed in the cochlear nucleus labels radial fibres (RF) of spiral ganglion neurons (SGN) that project to Atoh1/GFP$^+$/Myo7a$^+$ cells (arrow and arrowhead). Inset, basal regions of indicated cells. **b**, Neurofilament-positive (NF$^+$) processes associate with Myo7a$^+$ inner and outer hair cells (apical). NF$^+$ processes terminated at the base of an Atoh1/GFP$^+$/Myo7a$^+$ cell cluster as seen in the 90° rotation of the apical three-dimensional reconstruction. **c**, Ctbp2 (red) was localized to the basolateral domain at P35. **d**, The Ctbp2 (red) signal in **c** that was contained within the Myo7a$^+$ hair-cell cytoplasm. Scale bars, 50 μm (**a**); 20 μm (**a**, inset); 10 μm (**b**–**d**).

from the organ of Corti towards the modiolus were occasionally observed (Fig. 2c, e, and Supplementary Fig. 4f).

The total number of Myo7a$^+$ cells (that is, putative hair cells, whether transfected or not) in the cochlear base was increased 1.8-fold, and 50% of these cells were Atoh1/GFP$^+$ (Supplementary Table 2). In addition, we detected a 2.2-fold increase in putative inner hair cells (IHCs) and a 1.5-fold increase in putative outer hair cells (OHCs) in the base. Consequently, the ratio of OHC to IHC shifted from 3.4:1 to 2.4:1 as a result of the disproportionate increase in IHCs. Little is known about inner and outer hair-cell fate specification and whether the cues responsible are temporally or spatially regulated (or both). *In utero* gene transfer of *Atoh1* at different stages of otic vesicle development may provide insights into the molecular mechanisms governing hair-cell fate specification.

There were fewer supernumerary Myo7a$^+$ cells generated in the midbase (1.2-fold increase; 40% Atoh1/GFP$^+$), and none in the apex (Supplementary Table 2). The observed base-to-apex gradation in the abundance of supernumerary Myo7a$^+$ cells is inversely associated with the gradient of cell cycle exit in the organ of Corti, in which apical progenitors exit from the cell cycle first and basal progenitors exit last[5,23]. We propose that the large increase in supernumerary hair cells in the base results from clonal expansion of *Atoh1*-transfected progenitors.

To test whether the supernumerary Atoh1/GFP$^+$/Myo7a$^+$ cells attracted nerve fibres, we labelled neurons in the cochlear nucleus in retrograde fashion[24] and analysed the distribution of labelled processes to the organ of Corti. A cluster of four Atoh1/GFP$^+$/Myo7a$^+$ cells in the OHC region attracted a cluster of fibres that labelled with NeuroVue Red[25] (Fig. 3a, arrowhead) whose density was enriched at the base of the cluster (Fig. 3a, inset, arrowhead). An isolated Atoh1/GFP$^+$ cell in the same organ of Corti (Fig. 3a, arrow) attracted a refined NeuroVue-positive fibre to its base (Fig. 3a, inset, arrow). To confirm that the labelled fibres were neuronal processes, we analysed neurofilament distribution in Atoh1/GFP-transfected cochleae. We detected neurofilament-positive processes terminating at the base of both untransfected Myo7a$^+$ cells and Atoh1/GFP$^+$/Myo7a$^+$ cells (Fig. 3b and Supplementary Video 3). We next sought to determine whether the Atoh1/GFP-transfected cells expressed carboxy-terminal binding protein 2 (Ctbp2), a marker of the hair-cell ribbon synapse, at postnatal day 35 (P35). We detected discrete foci of Ctbp2 in the basolateral domain in 20 of 39 Atoh1/GFP$^+$ cells from two cochleae (Fig. 3c, d, and Supplementary Video 4). These data suggest that Atoh1/GFP$^+$ cells engage a synaptogenic program and are innervated by neurons that associate with the cochlear nucleus.

Hair cells convert mechanical stimuli into electrical impulses and have a distinct set of biophysical properties. To characterize the identity of Atoh1/GFP$^+$ cells further, we interrogated their electrophysiological properties at P4–P6. Cells were selected and grouped on the basis of whether they were GFP$^+$ or GFP$^-$ and whether they had the morphological phenotype of an IHC or an OHC. Morphology was based on the shape of the hair bundle viewed by differential interference contrast microscopy, with OHCs having the classical V shape and the IHCs being straight. OHCs were located on the strial side of pillar cells and IHCs were located on the modiolar side (Fig. 2a). Supernumerary IHCs and OHCs were investigated. The ectopic cells displaced towards the modiolus were structurally unstable and could not be patch-clamped (Supplementary Fig. 4f). Comparisons were made both between cell types and within types for GFP$^+$ and GFP$^-$ cells. Where no differences were observed, data were pooled.

No difference was observed in zero-current potential between GFP$^+$ and GFP$^-$ hair cells, or between IHCs and OHCs, with values of -63 ± 13 mV ($n = 5$) and -49 ± 15 mV ($n = 7$), respectively. Although membrane capacitance was not different between GFP$^+$ and GFP$^-$ cells, capacitance was different ($P < 0.001$, two-tailed *t*-test) between IHCs and OHCs, with values of 7.3 ± 1.3 pF ($n = 11$) and 4.7 ± 1.5 pF ($n = 24$), respectively. These results indicate that Atoh1/GFP-transfected and untransfected cells elaborated differential capacitances consistent with their identities as inner or outer hair cells.

Mechanotransduction was investigated by stimulating hair bundles with a piezoelectrically driven glass fibre (Fig. 4a)[26]. We found that hair cells induced by *Atoh1* misexpression had the same range of current amplitudes, sensitivity and adaptation as hair cells not expressing GFP. Figure 4 shows this analysis for GFP$^+$ and GFP$^-$ OHCs. Current amplitudes varied considerably as expected, in part because of the developmental age at which they were measured[26]. Peak currents of 152 ± 107 pA ($n = 15$) and 200 ± 177 pA ($n = 7$) were measured for GFP$^+$ and GFP$^-$ cells, respectively (Fig. 4b, c). Normalized current–displacement plots (Fig. 4d) were fitted with a single Boltzmann function, and no differences were found either in half activating displacement or in sensitivity, with values of 0.54 ± 0.06 μm and 0.53 ± 0.02 μm and slopes of 0.17 ± 0.05 μm^{-1} and 0.17 ± 0.03 μm^{-1} ($r^2 = 0.99$ for both) for GFP$^+$ and GFP$^-$ cells, respectively. Adaptation kinetics were compared by fitting the response to a positive stimulus that elicited less than 50% of the maximal response with the equation for a double exponential[27]. Fast and slow time constants did not differ between groups and gave values of 0.81 ± 0.48 ms ($n = 6$) and 0.39 ± 0.3 ms ($n = 4$) for the fast time constant and 8 ± 10 ms and 8 ± 6 ms for the slow time constant for GFP$^+$ and GFP$^-$ OHCs, respectively. These results suggest that mechanotransduction follows a normal developmental progression in Atoh1/GFP$^+$ hair cells[26].

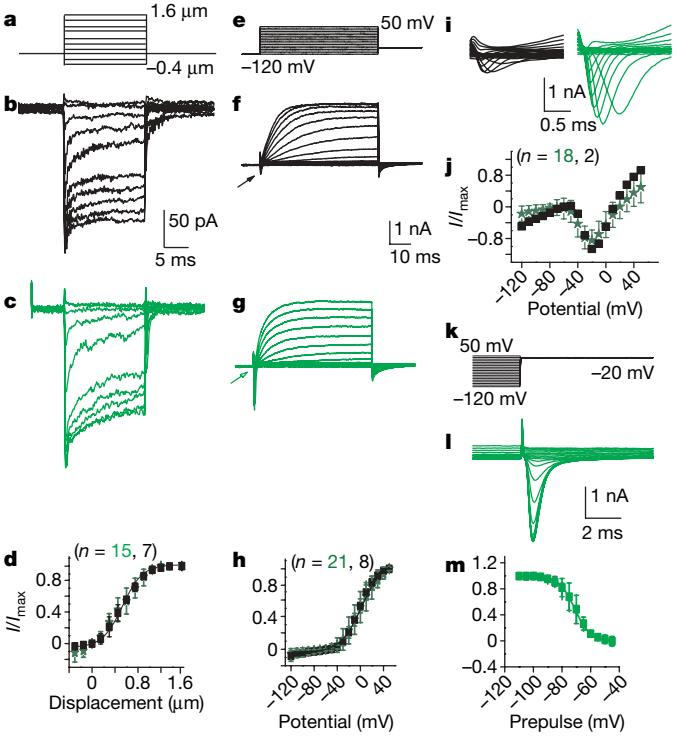


Figure 4 | Atoh1/GFP$^+$ cells mechanotransduce and elaborate age-appropriate basolateral conductances. **a**, Biophysical response properties of control (black) and GFP$^+$ (green) hair cells. **b**, **c**, Mechanically evoked currents from a holding potential of −84 mV elicited from the stimulus in **a**. **d**, Normalized peak current against displacement; the solid line is a Boltzmann fit. **f**, **g**, Current responses to the voltage-clamp stimuli in **e** for control (**f**) and GFP$^+$ (**g**) hair cells. **h**, Steady-state current–voltage plot from the data in **f** and **g**; the solid line is a Boltzmann fit. **i**, Expanded current traces from the regions indicated by the arrows in **f** and **g**. **j**, Plot of normalized peak current against potential for inward current. **k**, Stimulus protocol to probe inactivation that elicited currents in **l**. **m**, Prepulse potential against normalized peak inward current; the solid line is a Boltzmann fit. Data are presented as means ± s.d., and two-tailed *t*-tests were used for statistical comparisons.

Basolateral conductances were also investigated (Fig. 4e–m). A voltage protocol (Fig. 4e) that first prepulsed the potential from −84 mV to −120 mV for 200 ms was used to determine whether inactivating conductances were present. Both inward and outward currents were observed (Fig. 4f, g, i). The outward currents did not vary in current amplitude or activation properties between GFP^+ and GFP^- cells; however, there was a difference ($P < 0.001$) between OHCs and IHCs in current amplitude, with OHCs having a peak current of 1,878 ± 653 pA ($n = 29$) and IHCs having a peak current of 2,790 ± 767 pA ($n = 14$). These results again demonstrate that Atoh1/GFP^+ cells develop outward current characteristics consistent with their terminal differentiation as inner or outer hair cells.

Rapidly activating, rapidly inactivating inward currents were observed in 90% of the GFP^+ OHCs but in only 30% of GFP^- cells (Fig. 4f, g, i–m). The trend was less apparent for IHCs, with 80% of GFP^+ and 66% of GFP^- IHCs having inward currents. The greatest difference in the inward current is shown in Fig. 4i; the GFP^+ OHCs had much larger current amplitudes than the GFP^- cells did, with peak currents of 1,998 ± 1,178 pA ($n = 19$) and 475 ± 100 pA ($n = 3$), respectively ($P < 0.05$). Persistent, strong expression of Atoh1/GFP at postnatal stages by the *EEF1A1* promoter may underlie the difference in current amplitude observed. No difference was found in activation properties between any of the groups. Inactivation was not different between GFP^+ and GFP^- cells, but again there was a difference in the half inactivation between GFP^+ IHCs and OHCs, with values of −78 ± 1 mV and −71 ± 1 mV, respectively ($P < 0.01$, $n = 4$ for each). No difference was found for the slope, with values of 6.0 ± 1 mV^{-1} and 5.5 ± 0.4 mV^{-1} for IHCs and OHCs, respectively. These data support the conclusion that misexpression of *Atoh1* leads to the production of functional sensory hair cells in the postnatal cochlea that elaborate electrophysiological specializations consistent with inner or outer hair-cell identity.

There is close interest in gene[10,11] and cell[28,29] replacement strategies to restore auditory and vestibular function in the diseased inner ear. Although *Atoh1* misexpression in cochlear cultures[7,8] and in the adult guinea pig inner ear[9,10] generates cells with some of the morphological and molecular characteristics of sensory hair cells, no electrophysiological analysis of the induced hair cells was conducted in previous studies. We microinjected an expression plasmid encoding Atoh1 into the nascent mouse otocyst and subsequently transfected otic epithelial progenitor cells by *in vivo* electroporation. The induced supernumerary hair cells express Myo7a, attract neurofilament-bearing processes and localize the ribbon synapse marker Ctbp2 to their basolateral domain. In addition, the induced hair cells mechanotransduce and possess basolateral currents consistent with their maturational stage. Our work establishes that supernumerary cells induced by electroporation-mediated gene transfer of *Atoh1* are functional sensory hair cells. The ability to conduct gain-of-function experiments in the developing mammalian inner ear by gene transfer *in utero* may permit the design and functional assessment of gene therapies aimed at ameliorating hearing loss and vestibular dysfunction in mice that model human deafness and balance disorders. This capability is a crucial first step in defining translational therapies to ameliorate the effects of inner-ear disease in humans.

METHODS SUMMARY

Experimental embryology. Timed pregnant dams were laparotomized to externalize the uterus. Four to six otocysts were injected and electroporated in each dam. Transfected inner ears were harvested at E12.5, E18.5, P4–P6 or P28–P35 and fixed in 4% paraformaldehyde in PBS before cryosectioning, nerve tract tracing, scanning electron microscopy or immunohistochemistry.

Nerve tract tracing. The E18.5 crania were hemisected sagittally, and a nylon filter coated with NeuroVue Red was inserted into the cochlear nucleus followed by incubation for 7 days at 37 °C in PBS/0.1% sodium azide. Nerve tracts were revealed by confocal microscopy in cochlear whole mounts immunostained for Myo7a.

Scanning electron microscopy. The OTOTO method was performed[30].

Video. Confocal stacks of immunostained cochleae were processed with Imaris software to generate three-dimensional images and video.

Cell quantification. The base, midbase and apex of transfected E18.5 cochleae immunostained for Myo7a were imaged by confocal microscopy. The mean number of hair cells, transfected hair cells and transfected supporting cells was counted per 100 μm field (mean ± s.e.m.), and the Mann–Whitney *U*-test was used for statistical comparisons.

Electrophysiology. Hair cells were patch-clamped with an axoclamp 200b amplifier coupled to a 1322 digidata A/D, D/A board driven by JClamp software. Soda-glass pipettes (resistances 1–3 MΩ) coated with ski wax were used to record from both GFP^+ and GFP^- cells. Mechanical stimulation was accomplished with a stiff glass probe attached to a piezoelectric stack. Stimuli were filtered at 20 kHz and differentially amplified through an attenuator to regulate stimulus amplitudes. Data were sampled at 40 kHz and filtered at 10 kHz; each stimulus is an average of eight samples. Junction potential and series resistance were corrected offline. No leak subtraction was applied. Data are presented as means ± s.d., and two-tailed *t*-tests were used for statistical comparisons.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 1 April; accepted 17 July 2008.
Published online 27 August 2008.

1. Vollrath, M., Kwan, K. & Corey, D. The micromachinery of mechanotransduction in hair cells. *Annu. Rev. Neurosci.* **30,** 339–365 (2007).
2. Grant, L. & Fuchs, P. A. Auditory transduction in the mouse. *Pflugers Arch.* **454,** 793–804 (2007).
3. Davis, A. in *Hearing Science and Hearing Disorders* (eds Lutman, M. & Haggard, M.) (Academic, 1993).
4. Bermingham, N. A. *et al. Math1*: an essential gene for the generation of inner ear hair cells. *Science* **284,** 1837–1841 (1999).
5. Chen, P., Johnson, J. E., Zoghbi, H. Y. & Segil, N. The role of Math1 in inner ear development: Uncoupling the establishment of the sensory primordium from hair cell fate determination. *Development* **129,** 2495–2505 (2002).
6. Jones, J. M. *et al.* Inhibitors of differentiation and DNA binding (Ids) regulate Math1 and hair cell formation during the development of the organ of Corti. *J. Neurosci.* **26,** 550–558 (2006).
7. Zheng, J. L. & Gao, W. Q. Overexpression of *Math1* induces robust production of extra hair cells in postnatal rat inner ears. *Nature Neurosci.* **3,** 580–586 (2000).
8. Woods, C., Montcouquiol, M. & Kelley, M. W. Math1 regulates development of the sensory epithelium in the mammalian cochlea. *Nature Neurosci.* **7,** 1310–1318 (2004).
9. Kawamoto, K. *et al. Math1* gene transfer generates new cochlear hair cells in mature guinea pigs *in vivo*. *J. Neurosci.* **23,** 4395–4400 (2003).
10. Izumikawa, M. *et al.* Auditory hair cell replacement and hearing improvement by *Atoh1* gene therapy in deaf mammals. *Nature Med.* **11,** 271–276 (2005).
11. Staecker, H., Praetorius, M., Baker, K. & Brough, D. E. Vestibular hair cell regeneration and restoration of balance function induced by math1 gene transfer. *Otol. Neurotol.* **28,** 223–231 (2007).
12. Knirsch, M. *et al.* Persistence of $Ca_v1.3$ Ca^{2+} channels in mature outer hair cells supports outer hair cell afferent signaling. *J. Neurosci.* **27,** 6442–6451 (2007).
13. Wan, L., Almers, W. & Chen, W. Two *ribeye* genes in teleosts: the role of Ribeye in ribbon formation and bipolar cell development. *J. Neurosci.* **25,** 941–949 (2005).
14. Brown, S. D., Hardisty-Hughes, R. E. & Mburu, P. Quiet as a mouse: dissecting the molecular and genetic basis of hearing. *Nature Rev. Genet.* **9,** 277–290 (2008).
15. Friedman, L. M., Dror, A. A. & Avraham, K. B. Mouse models to study inner ear development and hereditary hearing loss. *Int. J. Dev. Biol.* **51,** 609–631 (2007).
16. Gubbels, S., Woessner, D., Mitchell, J. & Brigande, J. in *Thirtieth Annual MidWinter Research Meeting of the Association for Research in Otolaryngology, 10–15 February 2007* (ed. Santi, P. A.) vol. 30 330 (Association for Research in Otolaryngology, 2007).
17. Kim, D. W. *et al.* Use of the human elongation factor 1α promoter as a versatile and efficient expression system. *Gene* **91,** 217–223 (1990).
18. Li, X. *et al.* Generation of destabilized green fluorescent protein as a transcription reporter. *J. Biol. Chem.* **273,** 34970–34975 (1998).
19. Rechsteiner, M. PEST sequences are signals for rapid intracellular proteolysis. *Semin. Cell Biol.* **1,** 433–440 (1990).
20. Sher, A. E. The embryonic and postnatal development of the inner ear of the mouse. *Acta Otolaryngol.*, Suppl. 285, 1–77 (1971).
21. Hasson, T. *et al.* Unconventional myosins in inner ear sensory epithelia. *J. Cell Biol.* **137,** 1287–1307 (1997).
22. Hasson, T. *et al.* Expression in cochlea and retina of myosin VIIa, the gene product defective in Usher syndrome type 1B. *Proc. Natl Acad. Sci. USA* **92,** 9815–9819 (1995).
23. Chen, P. & Segil, N. $p27^{Kip1}$ links cell proliferation to morphogenesis in the developing organ of Corti. *Development* **126,** 1581–1590 (1999).
24. Maklad, A. & Fritzsch, B. Partial segregation of posterior crista and saccular fibers to the nodulus and uvula of the cerebellum in mice, and its development. *Brain Res. Dev. Brain Res.* **140,** 223–236 (2003).
25. Fritzsch, B. *et al.* Diffusion and imaging properties of three new lipophilic tracers, NeuroVue Maroon, NeuroVue Red and NeuroVue Green and their use

for double and triple labeling of neuronal profile. *Brain Res. Bull.* **66,** 249–258 (2005).
26. Waguespack, J., Salles, F. T., Kachar, B. & Ricci, A. J. Stepwise morphological and functional maturation of mechanotransduction in rat outer hair cells. *J. Neurosci.* **27,** 13890–13902 (2007).
27. Wu, Y. C., Ricci, A. J. & Fettiplace, R. Two components of transducer adaptation in auditory hair cells. *J. Neurophysiol.* **82,** 2171–2181 (1999).
28. Li, H., Corrales, C. E., Edge, A. & Heller, S. Stem cells as therapy for hearing loss. *Trends Mol. Med.* **10,** 309–315 (2004).
29. Hu, Z. & Ulfendahl, M. Cell replacement therapy in the inner ear. *Stem Cells Dev.* **15,** 449–459 (2006).
30. Self, T. *et al.* Shaker-1 mutations reveal roles for myosin VIIA in both development and function of cochlear hair cells. *Development* **125,** 557–566 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank C. Bresee and J. Jungwirth for expert technical support; C. Cepko and G. Nolan for plasmids EF1α–GFP and BMN–IRES–GFP, respectively; A. Kiernan, D. Fekete, S. Heller, A. Nguyen-Huynh, C. Bresee and J. Jungwirth for critical comments that improved the manuscript; D. Trune, B. Fritzsch and N. Segil for helpful discussions; S. Griest for statistical analyses; and M. Campbell and S. Nigra for exceptional animal care. This study was supported by grants from the National Institute on Deafness and Other Communication Disorders (J.V.B. and A.J.R.), the McKnight Endowment Fund for Neuroscience (J.J.B.) and the American Otological Society (Research Training Fellowship to S.P.G.).

Author Contributions The project was conceived by J.V.B. Experiments were planned and performed by S.P.G. and D.W.W. with advice from J.V.B., and were analysed by S.P.G., D.W.W. and J.V.B. J.V.B. performed the experimental embryology. A.J.R. conducted the electrophysiology experiments and interpreted the results. J.C.M. acquired the scanning electron micrographs. J.V.B. and A.J.R. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.V.B. (brigande@ohsu.edu).

METHODS

Expression plasmid construction. To construct pEF1α–GFP we cut pBMN-IRES-GFP and pEF1α-GFP with EcoRI/BsrGI to excise IRES–GFP and GFP, respectively. IRES–GFP was then ligated to the pEF1α vector to form pEF1α–IRES–GFP. To construct EF1α–ZsGreen, we subcloned the *EEF1A1* promoter from pEF1α–GFP into pGEM with the use of SalI. A plasmid with *EEF1A1* in the correct orientation was then digested with SalI/EcoRI and the promoter fragment was ligated into pZsGreen1-DR (Clontech) linearized with XhoI/EcoRI. To construct pEF1α–Atoh1–IRES–GFP, we created the Gateway-enabled destination vector pEF1α–RfC.1–IRES–GFP with the Gateway Vector Conversion System (Invitrogen) by digesting pEF1α–IRES–GFP with EcoRI, blunting the ends, and inserting Reading Frame Cassette C (RfC.1). Atoh1 complementary DNA was obtained from the American Type Culture Collection (MGC-19141) and the open reading frame was amplified with the primers 5′-GGG-GACAAGTTTGTACAAAAAAGCAGGCTTAATGTCCCGCCTGCTGCAT-3′ and 5′-GGGGACCACTTTGTACAAGAAAGCTGGGTACTAACTGGCCTCA-TCAGA-3′. A BP recombination reaction was performed with *Atoh1* cDNA and the donor vector pDONR/Zeo, creating pENTR-Atoh1/Zeo. The subsequent LR recombination reaction between pENTR-Atoh1/Zeo and pEF1α–RfC.1–IRES–GFP generated the expression plasmid pEF1α–Atoh1–IRES–GFP.

Expression plasmid purification. Expression plasmids were prepared with the Qiagen HiSpeed Plasmid Maxi Kit with modification. The plasmid was filtered at 0.22 μm before precipitation with ethanol, then resuspended in sterile PBS at 3 μg μl^{-1} and stored at −20 °C. Crystalline fast green was added to a freshly thawed aliquot to aid visualization during microinjection.

Timed pregnant breeding. Noon on the day on which a vaginal plug was detected was designated embryonic day 0.5 (E0.5) of development. Our initial experiments to define efficacious electroporation parameters were conducted with CD1 mice because of their fecundity. However, CD1 mice show elevated auditory thresholds by the fourth postnatal week. We outcrossed CD1 females with C57BL/6NTac males to generate CD1/B6 mice, which have normal auditory brainstem responses up to P35. Electrophysiology was conducted on hair cells in the CD1/B6 postnatal organ of Corti.

Experimental embryology. Dams were anaesthetized with 7.2 μl per gram body weight of a solution containing 9 mg ml^{-1} Nembutal, 20.8 mg ml^{-1} magnesium sulphate heptahydrate, 40% propylene glycol and 10% ethanol. The bicornate uterus was externalized by ventral laparotomy. A soft-cable fibre-optic light guide was pressed lightly against the irrigated uterus to illuminate the rostral and caudal branches of the primary head vein between which the otic vesicle resides (see Supplementary Video 1). Microinjection pipettes were fabricated with an 18–24 μm outer diameter and a 20° bevel. The pipette was advanced through the uterus and extraembryonic membranes and into the fluid-filled lumen of the otocyst. Expression plasmid tinted with fast green tracking dye was injected under pressure (about seven to nine 10-ms pulses per otocyst at about 70–103 kPa) into the lumen of the vesicle with a Picospritzer III (source gas nitrogen, 99.9% purity). The plasmid-filled otocyst was centred in the circular 5-mm field of the tweezer-style electrode paddles by securing the uterus with a gentle grip from the cathode and anode disks. A train of five square-wave pulses (43 V per pulse at 50 ms per pulse and 950 ms interpulse delay) was delivered to drive the plasmid into ventral otic epithelial progenitors.

Immunohistochemistry. Embryos were harvested 24 h or 6–7 days after electroporation and fixed in 4% paraformaldehyde in PBS (PFA/PBS; pH 7.2–7.4) for 8–12 h at 4 °C, with gentle agitation. Postnatal mice were fixed by cardiac perfusion with PFA/PBS and inner ears were decalcified in disodium EDTA. For cryostat sections, the entire E12.5 head was cryoprotected in graded sucrose/PBS to 30%, embedded in OCT medium and serially sectioned at 12 μm in the coronal plane. For whole mounts, the cochlea was dissected free of the cartilaginous capsule and the lateral wall was removed. The cryostat sections or dissected cochleae were permeabilized and blocked in 0.2% saponin in blocking solution (PBS containing 1% BSA and 3% serum from the species in which the secondary antibody was generated). Myo7a and neurofilament antibodies were applied overnight at 4 °C with gentle agitation. Alexa-Fluor-conjugated secondary antibody was applied for 2 h at room temperature (22 °C). Phalloidin–Alexa-Fluor conjugates were applied for 30 min at room temperature in PBS. Sectioned cochleae and the base, midbase or apex of whole-mount cochleae were covered in VectaShield (Vector Laboratories) before epifluorescence or confocal analysis. Antibodies and labelling reagents used were as follows: Alexa-660-conjugated phalloidin (dilution 1:50; A22285; Molecular Probes), Myosin 7a (dilution 1:150; 25-6790; Proteus Biosciences, Inc.), neurofilament (dilution 1:1,000; ab10586; Abcam, Inc.), Alexa 568-conjugated goat anti-rabbit (dilution 1:300; A11036; Molecular Probes) and Cy5-conjugated goat anti-chicken (dilution 1:100; ab6569; Abcam, Inc.).

Electrophysiology. Electrophysiological investigations were performed on isolated organs of Corti between P4 and P6. Tissue was isolated as described previously[25]. Isolated tissue was placed into a coverslip-bottomed recording chamber and held in place with single strands of dental floss. The bath was perfused at a rate of 2.5 ml min^{-1} with a solution containing (in mM): 135 NaCl, 1 KCl, 10 HEPES, 1.5 $CaCl_2$, 2 $MgCl_2$, 6 glucose, 4 pyruvate, 2 ascorbate, 2 creatine; the pH was set at 7.4 and osmolality was maintained at 325 mosmol kg^{-1}. The chamber and tissue were viewed with an Olympus BX51 microscope with the use of a ×60 water-immersion lens and Nomarski optics (Olympus America); images were captured with a Hamamatsu C2400 camera. Hair cells were patch-clamped with an Axoclamp 200b amplifier (Molecular Devices) coupled to a 1322 Digidata A/D, D/A board (Molecular Devices) driven by JClamp software (SciSoft). Soda-glass (Garner Glass) pipettes (resistances 1–3 MΩ) coated with ski wax were used to record from both GFP^+ and GFP^- cells. The internal solution contained (in mM): 125 KCl, 10 HEPES, 5 Mg-ATP, 5 creatine phosphate, 1 EGTA, 2 ascorbate; the pH was balanced to 7.2 and the osmolality was kept at 295 mosmol kg^{-1}. Mechanical stimulation was accomplished with a stiff glass probe attached to a piezoelectric stack (Physik Instrumente). Stimuli were filtered at 20 kHz (901P filter; Frequency Devices) and differentially amplified through a PA5 attenuator (Tucker Davis) to regulate stimulus amplitudes. Data were sampled at 40 kHz, filtered at 10 kHz; each stimulus is an average of eight samples. Junction potential and series resistance were corrected offline.

LETTERS

FcRn-mediated antibody transport across epithelial cells revealed by electron tomography

Wanzhong He[1,3], Mark S. Ladinsky[4], Kathryn E. Huey-Tubman[1,2], Grant J. Jensen[1], J. Richard McIntosh[4] & Pamela J. Björkman[1,2]

The neonatal Fc receptor (FcRn) transports maternal IgG across epithelial barriers[1,2], thereby providing the fetus or newborn with humoral immunity before its immune system is fully functional. In newborn rats, FcRn transfers IgG from milk to blood by apical-to-basolateral transcytosis across intestinal epithelial cells. The pH difference between the apical (pH 6.0–6.5) and basolateral (pH 7.4) sides of intestinal epithelial cells facilitates the efficient unidirectional transport of IgG, because FcRn binds IgG at pH 6.0–6.5 but not at pH 7 or more[1,2]. As milk passes through the neonatal intestine, maternal IgG is removed by FcRn-expressing cells in the proximal small intestine (duodenum and jejunum); remaining proteins are absorbed and degraded by FcRn-negative cells in the distal small intestine (ileum)[3–6]. Here we use electron tomography to make jejunal transcytosis visible directly in space and time, developing new labelling and detection methods to map individual nanogold-labelled Fc within transport vesicles[7] and simultaneously to characterize these vesicles by immunolabelling. Combining electron tomography with a non-perturbing endocytic label allowed us to conclusively identify receptor-bound ligands, resolve interconnecting vesicles, determine whether a vesicle was microtubule-associated, and accurately trace FcRn-mediated transport of IgG. Our results present a complex picture in which Fc moves through networks of entangled tubular and irregular vesicles, only some of which are microtubule-associated, as it migrates to the basolateral surface. New features of transcytosis are elucidated, including transport involving multivesicular body inner vesicles/tubules and exocytosis through clathrin-coated pits. Markers for early, late and recycling endosomes each labelled vesicles in different and overlapping morphological classes, revealing spatial complexity in endo-lysosomal trafficking.

To prevent ligand misdirection caused by a bulky label, we covalently attached small (1.4-nm) Nanogold to IgG-Fc (Au-Fc) at a site distant from where FcRn binds[7]. To avoid attaching more than 1 ligand per gold particle, which could erroneously prolong release through avidity, we used monofunctional Nanogold and purified Au-Fc by sizing and FcRn-affinity chromatography[7]. For steady-state experiments, Au-Fc was fed to neonatal rats, rather than incubated with excised intestines, which causes morphological changes[5]. The concentration of ingested Au-Fc was roughly equal to IgG in rat milk, because the higher concentrations used previously[5] saturated FcRn, resulting in degradation of excess IgG[8]. Intestinal samples were prepared for electron tomography by high-pressure freezing, freeze-substitution fixation (HPF/FSF), the most accurate method for preserving dynamic trafficking events and ultrastructure[9], and we developed methods to enlarge endocytosed Nanogold during FSF[7,10].

Internal controls verified that enlarged gold accurately marked transported Au-Fc: first, gold was in physiologically relevant locations (apical surface, tubulovesicular compartments in proximal cells, and inside degradative compartments in distal cells), but not in nuclei, mitochondria, the ER or Golgi (Figs 1–5 and Supplementary Figs 1–7); second, at least 98% of particles in proximal cells were 6–7 nm from a membrane (Supplementary Table 1), which is consistent with Au-Fc bound to FcRn; and third, Au-Fc, but not Au-dextran, was enhanced in proximal (FcRn-positive) cells, whereas both were enhanced in distal (FcRn-negative) cells, reflecting receptor-mediated and fluid-phase uptake in the proximal and distal intestines, respectively (Fig. 1a and Supplementary Fig. 7).

More than 50 tomograms, each about 1.8 μm^3, were recorded from jejunal cells from Au-Fc-fed neonatal rats (steady-state experiments; Supplementary Table 1). For kinetic analysis, ligated intestinal lumens were incubated with Au-Fc (more than 50 pulse or pulse–chase tomograms or projections; Supplementary Table 3). We defined three jejunal cell regions (Fig. 1b), namely region 1, consisting of the microvilli and terminal web[11]; region 2, between the terminal web and nucleus; and region 3, the lateral intercellular space (LIS) (basolateral membrane) and nearby cytoplasmic regions (Supplementary Figs 3 and 4). Region 3 was considered separately from region 2 because vesicles near the LIS participate in exocytosis/endocytosis. Within these regions, we classified gold-containing features into categories (Supplementary Table 1 and Supplementary Fig. 8): clathrin-coated pits at the apical/basolateral membranes, regular ~60 nm diameter tubular vesicles (RTVs—uniform diameters; variable lengths), coated buds/tips on RTVs, coated and uncoated spherical vesicles, irregular >70 nm diameter tubular vesicles (ITVs—variable diameters and lengths), irregular non-tubular vesicles (INTVs), coated/uncoated bulbs in ITVs/INTVs, multivesicular bodies (MVBs), and MVB inner vesicles, protrusions and tubules. Compartments that contained enlarged Au-Fc were further characterized by immunolabelling with antibodies against early (EEA1 and Rab5), late (Rab7 and Rab9) and recycling (Rab11) endosomes[12]. RTVs labelled with EEA1 and Rab5, MVBs with Rab5, and ITVs and INTVs with all five markers, although mainly Rab5, Rab9 and Rab11 (Fig. 1h, i, Supplementary Fig. 9 and Supplementary Table 2). These results demonstrated the morphological complexity of the endo-lysosomal system, revealing that 'early', 'late' and 'recycling' endosomes, as defined by their expression of markers, do not represent single categories of vesicles.

In region 1 we found Au-Fc on microvillar surfaces (suggesting receptor-mediated uptake at acidic pH), and in 60–120-nm clathrin-coated pits (Fig. 1c–g, Supplementary Fig. 8 and Supplementary Movie 1). Most Au-Fc endocytosis involved coated pits at the base

[1]Division of Biology 114-96 and [2]Howard Hughes Medical Institute, California Institute of Technology, 1200 East California Boulevard, Pasadena, California 91125, USA. [3]Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117543. [4]Boulder Laboratory for 3D Electron Microscopy of Cells, Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, Colorado 80309, USA.

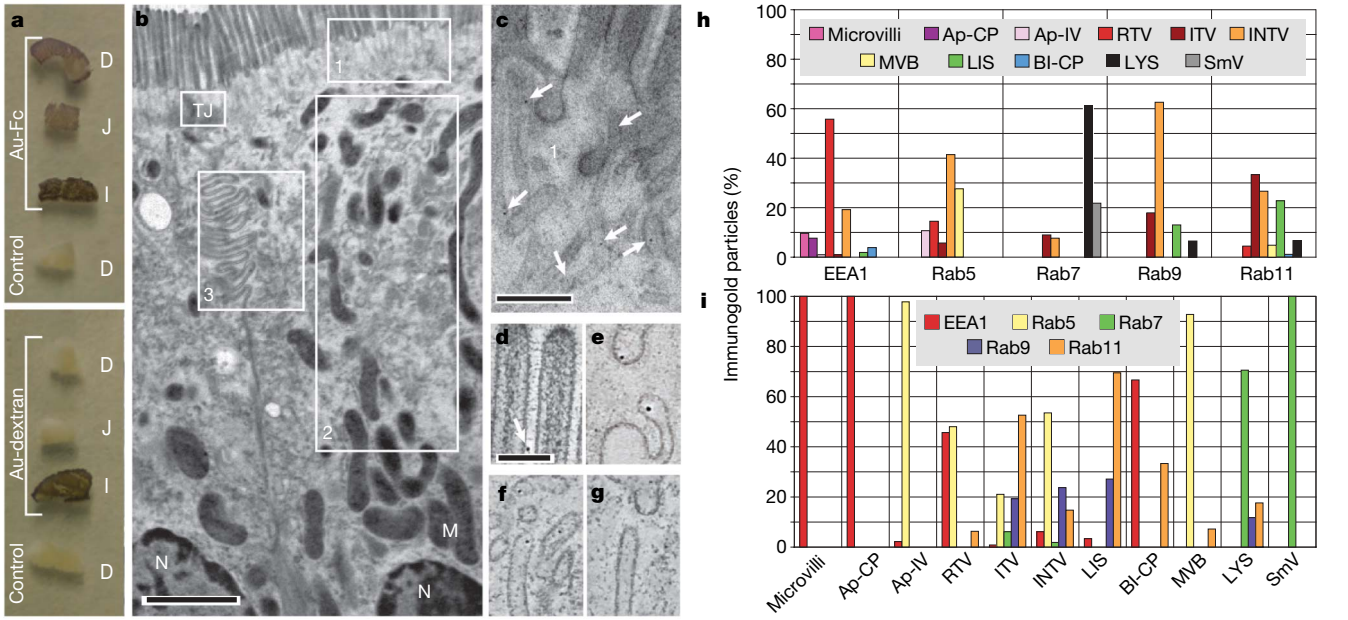


Figure 1 | Au-Fc uptake in intestinal cells. **a**, Chemically fixed and gold-enhanced intestinal samples from the duodenum (D), jejunum (J) and ileum (I) of Au-Fc-fed, Au-dextran-fed and control rats (Supplementary Fig. 7 shows corresponding projections). **b**, **c**, Projections of jejunal sections showing regions 1–3 (**b**; 180-nm section) or 1 (**c**; 150-nm section); see also Supplementary Figs 3–6. TJ, tight junction; N, nucleus; M, mitochondrion. White arrows indicate Au-Fc in RTVs. **d**–**g**, Tomographic slices showing Au-Fc in region 1 on microvilli (**d**), a coated pit and an apical irregular vesicle (**e**), RTVs and apical irregular vesicles (**f**) and RTV proximal to a coated pit (**g**). **h**, **i**, Histograms constructed with the immunolabelling data in Supplementary Table 2. Ap-CP, apical coated pits and coated vesicles; Ap-IV, apical irregular vesicles; Bl-CP, basolateral coated pits and coated vesicles; LYS, lysosomes; SmV, 30–40-nm vesicles related to lysosomes. **h**, Percentage of each marker found in each structure. **i**, Percentage of each structure that was positive for each marker. Scale bars, 2 μm (**b**); 300 nm (**c**); 200 nm (**d**–**g**).

of microvilli; many were proximal to gold-containing RTVs (Fig. 1c, g, and Supplementary Figs 3 and 4). Pulse experiments (Supplementary Table 3) suggested that the first transport steps involved a coated vesicle pinching off from a coated pit, uncoating, and fusing with an RTV or apical irregular vesicle.

In region 2, Au-Fc was in RTVs, ITVs, INTVs and MVBs. RTVs had nearly uniform diameters (Figs 1c and 2a, b, and Supplementary Fig. 8) but lengths between ~0.2 and 3.0 μm; they could branch, bend and loop (Supplementary Figs 3 and 4 and Supplementary Movie 2). Gold-containing spherical coated vesicles (Fig. 2c) were found throughout this region; their diameter of about 60 nm (Supplementary Fig. 8) suggested derivation from 60-nm coated tips/buds of RTVs (Fig. 2d), as found for transferrin receptor/transferrin trafficking[13]. Irregularly shaped vesicles, both non-tubular (INTVs) (Fig. 3a–c) and tubular (ITVs) (Figs 2e and 3d), also included gold. Similarly to RTVs, Au-Fc reached ITVs and INTVs within 5–7 min, but unlike RTVs, ITVs/INTVs were rarely in region 1, suggesting that they were populated after region 1 RTVs. Some

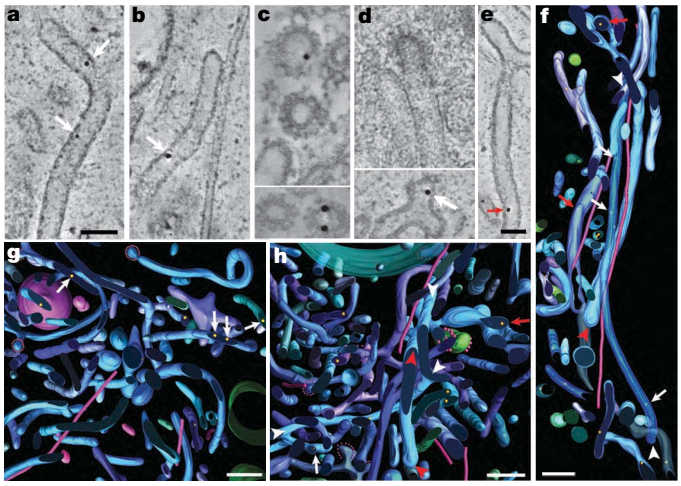


Figure 2 | Gold-containing tubular vesicles from jejunal regions 1 and 2. **a**, **b**, RTVs (associated with a microtubule in **b**). **c**, Small (less than 70-nm) coated vesicles. **d**, Small (less than 70-nm) coated buds. **e**, ITV. **f**–**h**, Models showing the cellular contexts of tubular vesicles (corresponding to Supplementary Movies 2 and 3, and Supplementary Figs 4b, 10a, b and 11a, b). the sample shown in **c** was chemically fixed; others were HPF/FSF. White arrowheads, RTVs; white arrows, Au-Fc in RTVs; red arrowheads, ITVs; red arrows, Au-Fc in ITVs. Colours for segmented models in all figures and movies are as follows: blue shades, tubular and irregular vesicles; pink straws, microtubules; violet, lysosomes; green with scarlet ribosomes, rough ER; bright green, mitochondria; purple and blue, LIS membranes; small red spheres around blue vesicles or buds, clathrin coats; gold spheres, Au-Fc. Scale bars, 100 nm (**a**–**d**, **e**); 200 nm (**f**–**h**).

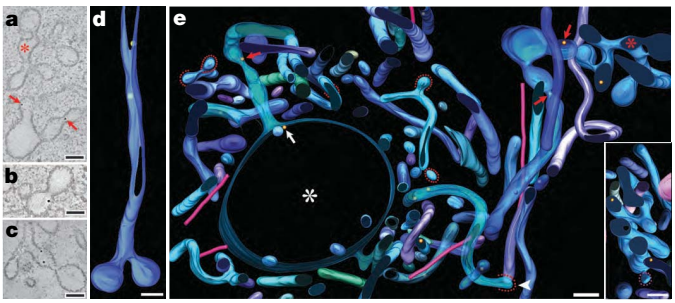


Figure 3 | Gold-containing irregular vesicles in jejunal regions 2 and 3. **a**–**c**, Tomographic slices. **a**, INTV (asterisk) and ITV (red arrows). **b**, INTV. **c**, INTV. **d**, **e**, Segmented models, coloured as in Fig. 2. **d**, ITV. **e**, The cellular context of irregular vesicles (see Supplementary Movie 4), including an MVB (white asterisk) with Au-Fc on the main membrane (white arrow) and in a tubular protrusion (red arrow). Top right of **e** corresponds to **a** (see red asterisks and arrows). Inset in **e** shows entangled irregular vesicles. Scale bars, 100 nm.

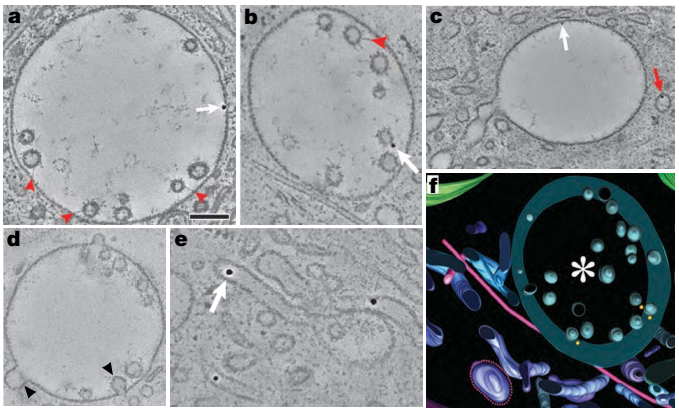


Figure 4 | Gold-containing MVBs in jejunal cells. Tomographic slices (**a**–**e**) and model (**f**, coloured as in Fig. 2). **a**, MVB showing fibrous extensions connecting inner vesicles with each other or the main membrane (red arrowheads) and Au-Fc on the main membrane (white arrow). **b**, MVB adjacent to a microtubule showing a fibrous extension (red arrowhead) and Au-Fc on an inner vesicle membrane (white arrow). **c**, Au-Fc on the main membrane of an MVB (white arrow) and in a spherical vesicle (red arrow). **d**, Invagination and protrusion (black arrowheads) in an MVB. **e**, Au-Fc in a tubular projection (white arrow). **f**, Model showing the MVB (white asterisk) in **b** (corresponding projection and larger model in Supplementary Figs 10c and 11c). Scale bar, 200 nm.

ITVs contained 70–130-nm bulbs (Fig. 3d) and/or small internal vesicles. INTVs included gold-containing 70–150-nm bulbs (Fig. 3c and Supplementary Movies 3 and 4). Coated bulbs on irregular vesicles were smaller (70–80 nm) than most uncoated bulbs (70–150 nm), but broader than the roughly 60-nm coated buds on RTVs (Supplementary Fig. 8). ITVs were often intertwined with INTVs and MVBs but not with RTVs (Figs 2f–h and 3e, Supplementary Fig. 11 and Supplementary Movies 2–4). Some gold-containing tubular vesicles were aligned with microtubules (Fig. 2b, f, and Supplementary Movies 2–4), suggesting microtubule-based motor movement, which is consistent with the disruption of FcRn-mediated transport by nocodazole[14].

Au-Fc appeared in MVBs in region 2 within 5 min of uptake and throughout a 25-min chase (Supplementary Table 3), always attached to a membrane (Fig. 4 and Supplementary Movies 4 and 5), suggesting binding to FcRn. In contrast, in the ileum, where IgG is degraded[5], gold was randomly distributed in MVB-like vesicles (Supplementary Fig. 2d, e). Jejunal MVBs appeared as predominantly Rab5-positive 0.3–1.2-μm spherical vesicles, sometimes associated with microtubules, which contained small (40–90 nm) inner vesicles (Figs 1h, i and 4, Supplementary Fig. 8 and Supplementary Table 2). As observed in dendritic and B cells[15], jejunal MVB inner vesicles were detached, not invaginating tubules open to cytoplasm (Supplementary Movies 4 and 5). Au-Fc was associated with the inner surface of the main MVB membrane and the outer surfaces of inner vesicles. Fusion of an FcRn–Fc-containing vesicle with an MVB accounted for the former; formation of an inner vesicle from a gold-associated membrane accounted for the latter. Accordingly, we observed invaginations and protrusions in the main membranes of MVBs (Fig. 4d).

Some jejunal MVBs contained tubular protrusions 50–130 nm in diameter (Supplementary Fig. 8), often resembling ITVs. Gold was found along the length and tips of these tubules, suggesting a reorganization of jejunal MVBs similar to that seen in dendritic cells when vesicles derived from MVB tubule tips carry peptide–MHC class II complexes to the cell surface[16]. A similar role for MVBs and

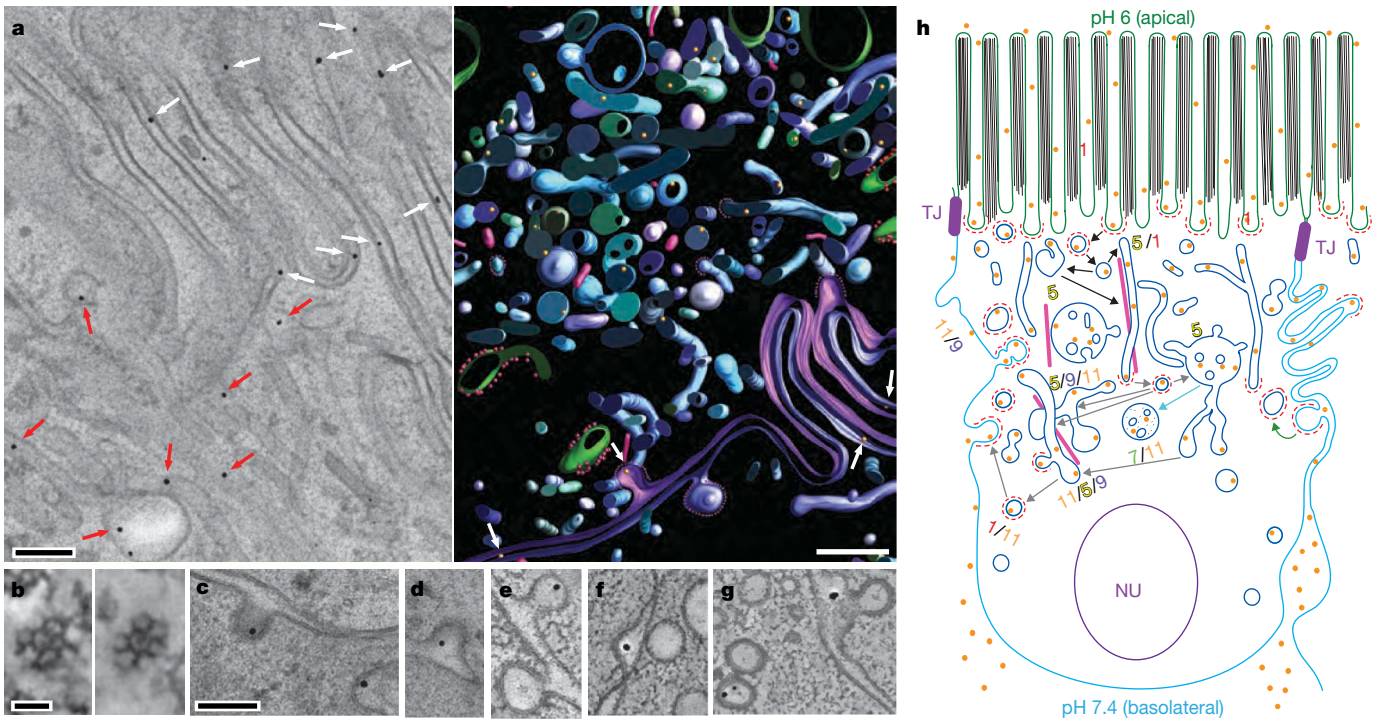


Figure 5 | Jejunal LIS and schematic pathways. **a**, LIS and vicinity as projection (left) and segmented model (right; coloured as in Fig. 2). White arrows, Au-Fc in LIS; red arrows, Au-Fc in irregular vesicles. **b**, Hexagonal (left) and heptagonal (right) clathrin lattices (chemically fixed samples). **c**, Au-Fc in a coated pit (projection image). **d**, Au-Fc in a coated pit (projection image). **e**, Au-Fc in a coated pit. **f**, Au-Fc in the LIS. **g**, Au-Fc in the LIS near a coated pit. **h**, Diagram of FcRn-mediated transcytosis. TJ, tight junctions; NU, nucleus; green, apical membrane; cyan, basolateral membrane; red dashes, clathrin; pink straws, microtubules; gold spheres, Au-Fc; black arrows, early trafficking steps; grey arrows, later trafficking steps; green arrow, endocytosis of FcRn-containing coated pit after ligand release; cyan arrow, excess Au-Fc transferred to late endosome (rarely observed). Multiple arrows indicate parallel pathways. Vesicles are shown with schematic models (Supplementary Fig. 8) at a lower density and larger size than observed. Numbers beside schematic vesicles refer to endosomal markers (1, EEA1; 5, Rab5; 7, Rab7; 9, Rab9; 11, Rab11); the font size reflects relative amounts of label in each type of compartment (see Fig. 1i). Scale bars, 200 nm (**a**, **c**–**g**); 50 nm (**b**).

their protrusions in Au-Fc transcytosis implies that MVBs serve primarily transport, not degradative, processes during FcRn-mediated transport across the intestine. In support of this notion, jejunal MVBs exhibited hallmarks of early endosomes: they contained relatively few internal vesicles[17] (Fig. 4) and labelled with markers for early and recycling, but not late, endosomes (Fig. 1h, i, and Supplementary Table 2).

Region 3 was distinguished from regions 1 and 2 by the presence of fewer tubular vesicles (Fig. 5a), more irregular vesicles, and an abundance of clathrin (Supplementary Fig. 11d and Supplementary Movie 6). Apical to the nucleus, Au-Fc was in ITVs, INTVs, 70–140-nm spherical vesicles, and coated pits within the LIS membrane. FcRn-mediated transcytosis of Au-Fc was demonstrated by extracellular gold in narrow regions and pockets of the LIS (Fig. 5a and Supplementary Movies 6–8) (accounting for about 10% of Au-Fc; Supplementary Table 1) and below the cell (Supplementary Figs 1f and 6). The first Au-Fc reached the LIS about 10 min after uptake, with most appearing after 15–30 min (Supplementary Table 3). Extracellular Au-Fc and intracellular gold-containing vesicles were largely apical to the nucleus.

Clathrin was found coating gold-containing vesicles, bulbs and pits, and free in the cytosol (Supplementary Movies 6 and 7). In chemically fixed samples, clathrin triskelions clearly formed the characteristic polygonal lattice (Supplementary Movie 6) containing mostly hexagons with occasional pentagons and heptagons (Fig. 5b). If, as is commonly assumed, clathrin coats are shed before vesicle fusion[18–21], gold-containing coated pits at the basolateral membrane would represent snapshots of endocytosis, which cannot normally be distinguished from exocytosis by EM. However, FcRn-mediated IgG transcytosis across the neonatal intestine is one of the few systems in which receptor-mediated reuptake of exocytosed ligands is strongly disfavoured: the neutral or basic pH at the basolateral surface should not allow large-scale FcRn-mediated endocytosis of Au-Fc, which was present at a concentration that would not allow binding to FcRn at pH $\geq$ 7 (Supplementary Discussion). The simplest interpretation of the gold-containing coated pits at the LIS membrane is that most or all are partially uncoated vesicles undergoing exocytosis. In support of this, we found at least 100 gold-containing coated pits in the basolateral membrane (Supplementary Table 1) but saw no definitive examples of gold in an uncoated pit. Thus, if these coated pits were not exocytosing, we did not observe any obvious exocytic events even though we found released Au-Fc in the extracellular space. Partially uncoated vesicles, which could accumulate during the slow, steady-state uncoating occurring after an initial burst of uncoating[22], might allow their uncoated portion to fuse with the basolateral membrane.

The existence of partly fused and partly coated vesicles is compatible with either 'kiss and run'[23] or 'prolonged release'[24] exocytosis, in which a transient pore opens to release the vesicle's contents. However, FcRn-mediated exocytosis cannot completely conform to the 'kiss and run' release of small molecules from 0.3–3.5-nm pores[25], because IgG and Fc dimensions approach 10 nm. Accordingly, gold-containing clathrin-coated pits in the LIS membrane were 90–150 nm, typically larger than their apical counterparts (60–120 nm), where mainly endocytosis is expected (Figs 1 and 5, and Supplementary Fig. 8).

As occurs in other exocytosis systems[26], clathrin-mediated endocytosis at the basolateral surface probably retrieves FcRn and the original vesicle after the release of IgG. We therefore propose that clathrin is maintained on membrane invaginations undergoing exocytosis to facilitate receptor recovery so that FcRn can contribute to another round of apical-to-basolateral transcytosis. This may be a specialized adaptation of neonatal jejunal cells for their primary function—IgG transport[5]. Indeed, clathrin levels near the LIS were unusually high in neonatal jejunal cells (Supplementary Movie 6) but become noticeably lower after weaning[5]. Further studies will be required to address definitively whether partly uncoated vesicles can fuse and, if so, whether unique features (for example, specialized machinery and/or low levels of vesicle-uncoating enzymes) facilitate this type of fusion in neonatal jejunal cells.

Steady-state and kinetic experiments revealed parallel transport pathways as summarized in Fig. 5h. In initial steps, Au-Fc enters jejunal cells in apical coated pits, and derived coated vesicles probably fuse, presumably after uncoating, with apical RTVs and/or with irregular vesicles. Coated buds on RTVs could serve as exit points, pinching off to shuttle Au-Fc to MVBs, ITVs and INTVs in regions 2 and 3. Au-Fc in clathrin-rich cytoplasmic regions near the LIS then transits through clathrin-coated spherical or irregular vesicles before exocytosis from a coated pit, which could be endocytosed to recycle FcRn. As a general feature, most potential transport vesicles contained a few labelled Fc molecules instead of a seemingly more efficient system in which a specialized subset of vesicles containing many labelled Fc molecules moved purposefully (for example, along microtubule tracks) to the basolateral membrane. Rather, most gold-containing vesicles were not visibly associated with microtubules, and vesicle-associated microtubules were not arranged in any particular direction.

Vesicle morphologies and intracellular contexts were compatible with kinetic data. First, RTVs, early endosomes containing about 20% of Au-Fc (Supplementary Table 1), were mainly in regions 1 and 2 and were not entangled with ITVs or INTVs. RTVs probably move FcRn–Fc complexes by microtubule-based travel or by releasing vesicles derived from roughly 60-nm coated buds. Both ITVs and INTVs usually had larger (70–80-nm) coated bulbs, as well as still larger uncoated bulbs, and these vesicles formed intertwined networks near the LIS with MVB tubules and each other (Fig. 3e). ITVs and INTVs accounted for about 34% of Au-Fc (Supplementary Table 1), suggesting an important role for irregular vesicles in FcRn-mediated transcytosis. Morphological similarities suggested that MVB tubules were related to ITVs (perhaps as precursors) and that ITV/INTV bulbs were related to basolateral coated pits, which were generally larger than their apical counterparts, suggesting a functional difference consistent with involvement in basolateral exocytosis.

Common steps between FcRn-mediated transcytosis and other intracellular trafficking include transit through acidic vesicles, exocytosis, and ligand release. Thus, our studies of FcRn transcytosis in epithelial cells are also relevant to understanding the role of FcRn in protecting IgG from degradation in endothelial cells[27] and other intracellular trafficking systems.

METHODS SUMMARY

Uptake of Au-Fc and Au-dextran. Monomaleimido-Nanogold (1.4 nm) was attached to one or more reduced cysteine residues in the rat Fc hinge, and *sulpho-N*-hydroxysuccinimido (sulpho-NHS)-Nanogold was attached to dextran primary amines as described[7]. Labelled molecules were purified by FcRn affinity chromatography and/or gel-filtration chromatography. For steady-state and immunolabelling experiments, 11–13-day-old suckling rats were fed with three 100-μl doses of Au-Fc or Au-dextran (about 3 μM), after which the small intestine was harvested for chemical or cryofixation. For kinetic experiments, ligated intestines were incubated with 3 μM Au-Fc (pulse experiments), followed by unlabelled Fc or IgG (pulse–chase experiments).

Gold enlargement procedures. Intestinal segments were chemically fixed, treated with GoldEnhance–EM 2113 (Nanoprobes, Inc.) and fixed/stained with OsO_4 and uranyl acetate as described[7]. Traditional pre-embedding enhancement protocols are incompatible with FSF, therefore we developed the methods for enlarging small gold clusters in the cold organic solvents used during FSF[7,10]. After HPF of intestinal segments, we used a three-step enlarging protocol involving silver enhancement to enlarge (to at most 8 nm) the Nanogold slightly[10]; coating the silver shell with gold to make it impervious to osmium; and enlargement to 10–16 nm by using gold enhancement[7]. Samples were then treated with OsO_4 and uranyl acetate and warmed to room temperature (about 22 °C).

Microscopy and three-dimensional modelling. HPF/FSF or chemically fixed samples were infiltrated with resin, polymerized, cut into 120–200-nm sections and examined with an FEI T12 transmission electron microscope operating at 120 kV. For tomography, dual-axis tilt series were collected at ×6,500 from 120–200-nm sections at 700 nm underfocus. Tomograms were computed for each tilt axis by using enlarged golds as alignment markers, and then aligned and

combined to form a dual-axis tomogram using IMOD[28]. Unless otherwise indicated, tomographic slices were 1.6 nm thick. Tomographic reconstructions were interpreted and modelled by manual segmentation with IMOD[28].

Immunolabelling. Immunolabelling was performed on gold-enhanced intestinal samples with a modification of the Tokuyasu method for immunolabelling cryosections[29].

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 14 April; accepted 14 July 2008.

1. Rodewald, R. & Kraehenbuhl, J.-P. Receptor-mediated transport of IgG. *J. Cell Biol.* **99,** S159–S164 (1984).
2. Simister, N. E. & Mostov, K. E. An Fc receptor structurally related to MHC class I antigens. *Nature* **337,** 184–187 (1989).
3. Rodewald, R. Selective antibody transport in the proximal small intestine of the neonatal rat. *J. Cell Biol.* **45,** 635–640 (1970).
4. Jones, E. A. & Waldmann, T. A. The mechanism of intestinal uptake and transcellular transport of IgG in the neonatal rat. *J. Clin. Invest.* **51,** 2916–2927 (1972).
5. Rodewald, R. Intestinal transport of antibodies in the newborn rat. *J. Cell Biol.* **58,** 189–211 (1973).
6. Rodewald, R. Distribution of immunoglobulin G receptors in the small intestine of the young rat. *J. Cell Biol.* **85,** 18–32 (1980).
7. He, W. *et al.* A freeze substitution fixation-based gold enlarging technique for EM studies of endocytosed Nanogold-labeled molecules. *J. Struct. Biol.* **160,** 103–111 (2007).
8. Benlounes, N. *et al.* Intestinal transport and processing of immunoglobulin G in the neonatal and adult rat. *Biol. Neonate* **67,** 254–263 (1995).
9. McIntosh, J. R., Nicastro, D. & Mastronarde, D. N. New views of cells in 3D: an introduction to electron tomography. *Trends Cell Biol.* **15,** 43–51 (2005).
10. Morphew, M., He, W., Bjorkman, P. J. & McIntosh, J. R. Silver enhancement of nanogold particles during freeze substitution fixation for electron microscopy. *J. Microsc.* **230,** 263–267 (2008).
11. Hull, B. E. & Staehelin, L. A. The terminal web. A reevaluation of its structure and function. *J. Cell Biol.* **81,** 67–82 (1979).
12. Schwartz, S. L., Cao, C., Pylypenko, O., Rak, A. & Wandinger-Ness, A. Rab GTPases at a glance. *J. Cell Sci.* **120,** 3905–3910 (2007).
13. Stoorvogel, W., Oorschot, V. & Geuze, H. J. A novel class of clathrin-coated vesicles budding from endosomes. *J. Cell Biol.* **132,** 21–33 (1996).
14. McCarthy, K. M., Yoong, Y. & Simister, N. E. Bidirectional transcytosis of IgG by the rat neonatal Fc receptor expressed in a rat kidney cell line: a system to study protein transport across epithelia. *J. Cell Sci.* **113,** 1277–1285 (2000).
15. Murk, J. L. A. N. *et al.* Endosomal compartmentalization in three dimensions: Implications for membrane fusion. *Proc. Natl Acad. Sci. USA* **100,** 13332–13337 (2003).
16. Kleijmeer, M. *et al.* Reorganization of multivesicular bodies regulates MHC class II antigen presentation by dendritic cells. *J. Cell Biol.* **155,** 53–63 (2001).
17. Mari, M. *et al.* SNX1 defines an early endosomal recycling exit for sortilin and mannose 6-phosphate receptors. *Traffic* **9,** 380–393 (2008).
18. Altstiel, L. & Branton, D. Fusion of coated vesicles with lysosomes: measurement with a fluorescence assay. *Cell* **32,** 921–929 (1983).
19. Brodsky, F. M., Chen, C. Y., Knuehl, C., Towler, M. C. & Wakeham, D. E. Biological basket weaving: formation and function of clathrin-coated vesicles. *Annu. Rev. Cell Dev. Biol.* **17,** 517–568 (2001).
20. Conner, S. D. & Schmid, S. L. Regulated portals of entry into the cell. *Nature* **422,** 37–44 (2003).
21. Bonifacino, J. S. & Glick, B. S. The mechanisms of vesicle budding and fusion. *Cell* **116,** 153–166 (2004).
22. Greene, L. E. & Eisenberg, E. Dissociation of clathrin from coated vesicles by the uncoating ATPase. *J. Biol. Chem.* **265,** 6682–6687 (1990).
23. Fesce, R., Grohovaz, F., Valtorta, F. & Meldolesi, J. Neurotransmitter release: fusion or 'kiss-and-run'? *Trends Cell Biol.* **4,** 1–4 (1994).
24. Ober, R. J., Martinez, C., Lai, X., Zhou, J. & Ward, E. S. Exocytosis of IgG as mediated by the receptor, FcRn: an analysis at the single-molecule level. *Proc. Natl Acad. Sci. USA* **101,** 11076–11081 (2004).
25. Sokac, A. M. & Bement, W. M. Kiss-and-coat and compartment mixing: coupling exocytosis to signal generation and local actin assembly. *Mol. Biol. Cell* **17,** 1495–1502 (2006).
26. Granseth, B., Odermatt, B., Royle, S. J. & Lagnado, L. Clathrin-mediated endocytosis is the dominant mechanism of vesicle retrieval at hippocampal synapses. *Neuron* **51,** 773–786 (2006).
27. Ghetie, V. & Ward, E. S. Multiple roles for the major histocompatibility complex class I-related FcRn. *Annu. Rev. Immunol.* **18,** 739–766 (2000).
28. Mastronarde, D. N. Dual-axis tomography: an approach with alignment methods that preserve resolution. *J. Struct. Biol.* **120,** 343–352 (1997).
29. Ladinsky, M. S. & Howell, K. E. Electron tomography of immunolabeled cryosections. *Methods Cell Biol.* **79,** 543–548 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Mastronarde for advice about setting up SerialEM on the Caltech microscopes and three-dimensional modelling; C. Kivork for preparing Au-Fc and Au-dextran; B. Tivol for assistance with electron microscopes; M. Murphy for help with schematic figures; M. Morphew for contributions to silver enhancement methods; Y. Nie, A. Feuerabendt, A. Kules, C. Luna, K. McKenzie, R. Sander and L. Zinn-Bjorkman for segmenting tomograms; J. Vielmetter for SPR studies of Au-Fc; and F. Brodsky, F. Maxfield, S. Schmid and D. Tesar for discussions. This work was supported by the National Institutes of Health (2 R37 AI041239-06A1 to P.J.B. and RR000592 to J.R.M.), a Max Planck Research Award (P.J.B.), gifts from the Gordon and Betty Moore Foundation and the Agouron Institute to support electron microscopy at Caltech, and National University of Singapore AcRF start-up funds (R-154-000-339-133 to W.H.).

Author Contributions W.H. and P.J.B. conceived the experiments. Electron microscopy data were collected by W.H. in collaboration with and in the microscopy laboratory established by G.J.J. at Caltech. W.H. configured SerialEM on the Caltech electron microscopes, prepared intestinal samples by chemical fixation or HPF/FSF, conceived and developed the gold enhancement procedures, and collected, processed, interpreted and modelled tomograms. W.H. and K.E.H.T conducted kinetic experiments. Immunolabelling and associated imaging was done by M.S.L. in Boulder with enhanced samples provided by W.H. and K.E.H.T. Initial phases of the project using transfected cells were conducted by P.J.B. in Boulder in collaboration with J.R.M. and the Boulder Laboratory for 3D Microscopy of Cells. All authors discussed and interpreted the results and edited the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to P.J.B. (bjorkman@caltech.edu).

METHODS

Uptake of gold-labelled ligands in neonatal rats. The preparation and characterization of gold-labelled Fc (Au-Fc) was described previously[7]. In brief, the hinge region of purified, reduced recombinant rat IgG2a Fc fragment (rFc) was labelled with 1.4-nm monomaleimido Nanogold (Nanoprobes, Inc.), which reacts specifically with reduced thiols, following the manufacturer's protocol for labelling IgG. Au-Fc was purified by gel-filtration and FcRn-affinity chromatography[7,30].

Red fluorescent dextran, a marker for fluid-phase uptake, was labelled with *sulfo*-NHS Nanogold (Nanoprobes, Inc.), which reacts with primary amines. *sulfo*-NHS Nanogold (30 nmol) was dissolved in 1.0 ml of deionized water, then added immediately to 0.3 mg of Texas Red Dextran (10 kDa) (Molecular Probes) and incubated for 2 h at room temperature. The reaction mixture was concentrated, then run over a Superdex S75 gel filtration column in 0.1 M $NaPO_4$ pH 7.5, 150 mM NaCl. Au-dextran was isolated as the slowest migrating red peak.

Sprague Dawley suckling rats 11–13 days old were fasted for about 3 h, and then fed with 100 µl of 2–3-µM Au-Fc or Au-dextran at time zero, then twice more at 40–60-min intervals as described[7]. After 90–150 min, a 4–5-cm segment of proximal small intestine, starting about 2 cm from the pylorus, and a 4–5-cm segment of the distal small intestine (ileum), located about 2 cm from the ileocecal valve, were removed for chemical fixation or HPF.

Kinetic experiments. Because it would be difficult, if not impossible, to synchronize the uptake of an ingested ligand in a live animal, we incubated ligated neonatal rat intestines with Au-Fc for fixed durations. For these experiments, 9–13-day-old rats were anaesthetized with 1–5% isoflurane, and an incision was made to expose the intestine. The proximal segment of the intestine was ligated with thread and then injected with 125–200 µl of 3 µM Au-Fc in 20 mM $NaPO_4$ pH 6.0. After incubation with Au-Fc for the indicated duration, the ligated intestine was either dissected and high-pressure frozen (pulse experiments) or washed with 2–5 volumes of 100 mM $NaPO_4$ pH 6.0, 2 mM $CaCl_2$, 1 mM $MgCl_2$, 0.5 mM $MgSO_4$ for further chase experiments. For some chase experiments, 200 µl of 5 µM unlabelled rFc was injected after the wash in pH 6.0 buffer and allowed to incubate for the additional durations indicated (Supplementary Table 3) before harvest and HPF. For other chase experiments, 125 µl of 300 µM human IgG (hIgG) was injected directly after the 5-min uptake period. The latest time point examined was 30 min after the initial incubation, in accordance with previous studies demonstrating that ingested IgG could reach the circulation of neonatal rats in 30 min (ref. 31). For experiments involving the processing of multiple segments, the labelling and chase times are listed as a range in Supplementary Table 3 to account for the time it took to harvest and freeze individual samples.

Gold enlargement techniques. The 1.4 nm Nanogold clusters were too small to be seen in tomograms recorded at the magnifications required to reconstruct an informative portion of a cell[7]. We therefore used enhancement techniques to enlarge the clusters. For chemically fixed samples, we used a modification of traditional gold enhancement techniques[32] that resulted in lower background and reduced auto-nucleation[7]. Chemical fixation, gold enhancement with GoldEnhance–EM 2113 (Nanoprobes, Inc.) and fixation/staining with OsO_4 and uranyl acetate were performed as described[7].

Traditional pre-embedding enhancement protocols to enlarge Nanogold are not compatible with FSF; we therefore developed the first reported enhancement methods for enlarging small gold clusters in the cold organic solvents used during FSF[7,10]. Intestinal segments were rapidly frozen with a Bal-Tec HPM 010 High Pressure Freezer (Bal-Tec AG) as described[7]. To avoid background resulting from spontaneous auto-nucleation, we used a three-step enlarging protocol in which silver enhancement was used to enlarge the Nanogold slightly (to not more than 8 nm; ref. 10); the silver shell was coated with gold to make it impervious to osmium; and particles were further enlarged to 10–16 nm by using gold enhancement[7]. After enhancement, samples were treated with OsO_4 and uranyl acetate and warmed to room temperature as described[7].

EM images of enhanced chemically fixed and HPF/FSF samples from Au-Fc-fed and control rats confirmed specific enhancement of Au-Fc with negligible background[7] (Fig. 1a and Supplementary Figs 1 and 7). The fact that gold particles of similar sizes were observed at all levels in the sections demonstrated that all particles were enlarged at about the same rate and that the tissue was completely penetrated by the enhancement procedure.

Microscopy and three-dimensional reconstructions. HPF/FSF or chemically fixed intestinal samples were infiltrated with resin, polymerized and sectioned[7,33]. Selected regions of 70–200-nm sections were examined with an FEI T12 transmission electron microscope operating at 120 kV, and projection images were recorded on a Gatan 894 2k × 2k charge-coupled device camera (Gatan Corporation). For tomography, 120–200-nm sections were coated on both sides with about 5 nm of evaporated carbon and images were recorded at 700 nm underfocus. Tilt series were digitally recorded at ×6,500 (1.573 nm per pixel) using the microscope control program SerialEM[34]. Dual-axis tilt series (±70°, 1.0° angular increments about each axis) were obtained with a high-tilt specimen holder (FEI) by removing the sample, rotating the grid through about 90°, and then reinserting the sample and collecting a second tilt series. Samples were preirradiated with about 10^5 electrons nm^{-2} to stabilize them before collecting images, and the cumulative electron dose during imaging was another roughly 10^5 electrons nm^{-2}. Tomograms were computed for each tilt axis using the enlarged gold particles as alignment markers; they were then aligned to each other and combined to form a dual-axis tomogram with the IMOD software package[28,35]. Tomographic slices shown in all figures are 1.6 nm thick unless otherwise indicated. For pulse and pulse–chase experiments, two-dimensional projections from thin sections (120 nm) were recorded on film with a Philips CM10 100-kV microscope.

Three-dimensional modelling. Tomographic reconstructions were interpreted and modelled with IMOD[28,35]. Tubular vesicles, microtubules and other organelles were segmented manually in tomograms from enhanced samples by tracing the portion of each object visible in single tomographic slices as described[36]. For clarity in some of the figures of three-dimensional scenes, the bottom and tops of a few vesicles were capped by adding several estimated contours.

Clathrin coats on vesicles and coated pits were easily identified in chemically fixed samples by the characteristic polygonal clathrin lattice (Supplementary Movie 6). Although the fixation and staining procedures required for gold enlargement in HPF/FSF samples did not preserve polygonal clathrin lattices, the spiky appearance of the membranes in equatorial sections allowed coated and uncoated pits and vesicles to be distinguished (Supplementary Movie 7).

Immunolabelling. Immunolabelling was performed on gold-enhanced intestinal samples with a modification of the Tokuyasu method for immunolabelling cryosections[37]. Jejunal samples were prepared from Au-Fc-fed neonatal rats as described[7]. Intestinal samples were injected with 4% paraformaldehyde in 100 mM phosphate pH 6.0 at 37 °C for 15 min, and then incubated in 4% paraformaldehyde in PBS pH 7.4 including 5% sucrose, first at room temperature for 1 h, and then on ice overnight. Au-Fc was enlarged by modifying the protocol described for chemically fixed samples[7] to include a final step in 4% paraformaldehyde and 5% sucrose. The samples were rinsed three times with fresh PBS, and then infiltrated with PBS containing 2.1 M sucrose for 4 h. The samples were affixed to aluminium pins (Ted Pella, Inc.) and quickly frozen in liquid nitrogen. Semi-thin cryosections (about 90 nm) were prepared as described[29], with an UltraCut UCT/FCS cryomicrotome (Leica Microsystems). Cryosections were transferred to a Formvar-coated, carbon-coated, 100-mesh copper/rhodium electron microscopy grid, warmed to room temperature and immunolabelled as described[29]. The following primary antibodies were used together with 10 nm colloidal gold-conjugated secondary antibodies (either goat anti-rabbit IgG or goat anti-mouse IgG from Ted Pella, Inc.): mouse monoclonal anti-EEA1 (5 µg ml^{-1}; BD Biosciences; catalogue no. 610457); rabbit polyclonal anti-Rab5 (2 µg ml^{-1}; Santa Cruz Biotechnology; catalogue no. sc-28570); rabbit polyclonal anti-Rab7 (H-50) (2 µg ml^{-1}; Santa Cruz Biotechnology; catalogue no. sc-10767), rabbit polyclonal anti-Rab9 (FL-201) (2 µg ml^{-1}; Santa Cruz Biotechnology; catalogue no. sc-28573) and mouse monoclonal anti-Rab11 (2.5 µg ml^{-1}; BD Biosciences; catalogue no. 610656).

30. Huber, A. H., Kelley, R. F., Gastinel, L. N. & Bjorkman, P. J. Crystallization and stoichiometry of binding of a complex between a rat intestinal Fc receptor and Fc. *J. Mol. Biol.* **230,** 1077–1083 (1993).
31. Halliday, R. The absorption of antibodies from immune sera by the gut of the young rat. *Proc. R. Soc. Lond. B* **143,** 408–413 (1955).
32. Danscher, G. Localization of gold in biological tissue. A photochemical method for light and electron microscopy. *Histochemistry* **71,** 1–16 (1981).
33. He, W., Cowin, P. & Stokes, D. L. Untangling desmosomal knots with electron tomography. *Science* **302,** 109–113 (2003).
34. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152,** 36–51 (2005).
35. Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer visualization of three-dimensional data using IMOD. *J. Struct. Biol.* **116,** 71–76 (1996).
36. Ladinsky, M. S., Mastronarde, D. N., McIntosh, J. R., Howell, K. E. & Staehelin, L. A. Golgi structure in three dimensions: functional insights from the normal rat kidney cell. *J. Cell Biol.* **144,** 1135–1149 (1999).
37. Tokuyasu, K. T. Application of cryoultramicrotomy to immunocytochemistry. *J. Microsc.* **143,** 139–149 (1986).

LETTERS

CDK8 is a colorectal cancer oncogene that regulates β-catenin activity

Ron Firestein[1,2,6,7], Adam J. Bass[1,3,6,7], So Young Kim[1,3,6,7], Ian F. Dunn[1,4,6,7], Serena J. Silver[7], Isil Guney[1,6,7], Ellen Freed[1], Azra H. Ligon[2], Natalie Vena[1], Shuji Ogino[1,2], Milan G. Chheda[1,5,7], Pablo Tamayo[7], Stephen Finn[2], Yashaswi Shrestha[1,6,7], Jesse S. Boehm[7], Supriya Jain[1], Emeric Bojarski[1], Craig Mermel[1,6,7], Jordi Barretina[1,6,7], Jennifer A. Chan[2,7], Jose Baselga[8], Josep Tabernero[8], David E. Root[7], Charles S. Fuchs[1], Massimo Loda[1,2], Ramesh A. Shivdasani[1,3], Matthew Meyerson[1,2,6,7] & William C. Hahn[1,3,6,7]

Aberrant activation of the canonical WNT/β-catenin pathway occurs in almost all colorectal cancers and contributes to their growth, invasion and survival[1,2]. Although dysregulated β-catenin activity drives colon tumorigenesis, further genetic perturbations are required to elaborate full malignant transformation[3]. To identify genes that both modulate β-catenin activity and are essential for colon cancer cell proliferation, we conducted two loss-of-function screens in human colon cancer cells and compared genes identified in these screens with an analysis of copy number alterations in colon cancer specimens. One of these genes, *CDK8*, which encodes a member of the mediator complex[4], is located at 13q12.13, a region of recurrent copy number gain in a substantial fraction of colon cancers. Here we show that the suppression of *CDK8* expression inhibits proliferation in colon cancer cells characterized by high levels of CDK8 and β-catenin hyperactivity. CDK8 kinase activity was necessary for β-catenin-driven transformation and for expression of several β-catenin transcriptional targets. Together these observations suggest that therapeutic interventions targeting CDK8 may confer a clinical benefit in β-catenin-driven malignancies.

To identify oncogenes that modulate β-catenin-dependent transcription and regulate colon cancer cell proliferation, we conducted two RNA interference (RNAi)-based loss-of-function screens. We engineered DLD-1 colon cancer cells, which harbour *APC* deletions

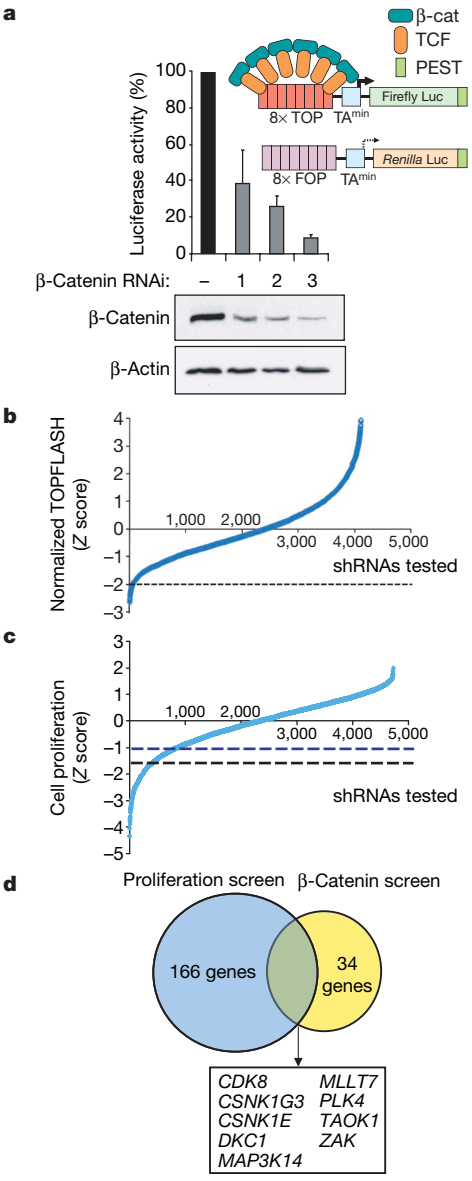


Figure 1 | RNAi screens to identify genes essential for colon cancer cell proliferation and β-catenin activity. **a**, Schematic of the DLD-1^{Rep} cell line showing the engineered 8× TOPFLASH (TOP) and 8× FOPFLASH (FOP) elements and relative TOP/FOP activity in the DLD-1^{Rep} cell line. TA, minimal TA promoter; PEST, domain rich in proline (P), glutamate (E), serine (S) and threonine (T) residues. **b**, Distribution curve showing *Z* scores representing β-catenin activity for all shRNAs tested in the DLD-1^{Rep} screen. shRNAs that reduced FOPFLASH levels to near background, or activated FOPFLASH more than 2 s.d. above the mean, were excluded (FOP ≤800 and FOP ≥2,600 luciferase units). Two out of five *CDK8*-specific shRNAs were excluded on this basis. shRNA that induced *Z* scores of >4 are not shown. The dashed line indicates the *Z* score cutoff for shRNAs scored as hits. **c**, Distribution curve showing *Z* scores representing cell proliferation for shRNAs tested in HCT116 cells. This screen contained shRNAs targeting 1,004 genes, and there was a 92% overlap between the screens in **b** and **c**. The blue and black dashed lines indicate a *Z* score cutoff for shRNAs scored as hits. **d**, Venn diagram representation of the nine genes that reduced both β-catenin activity and colon cancer cell proliferation.

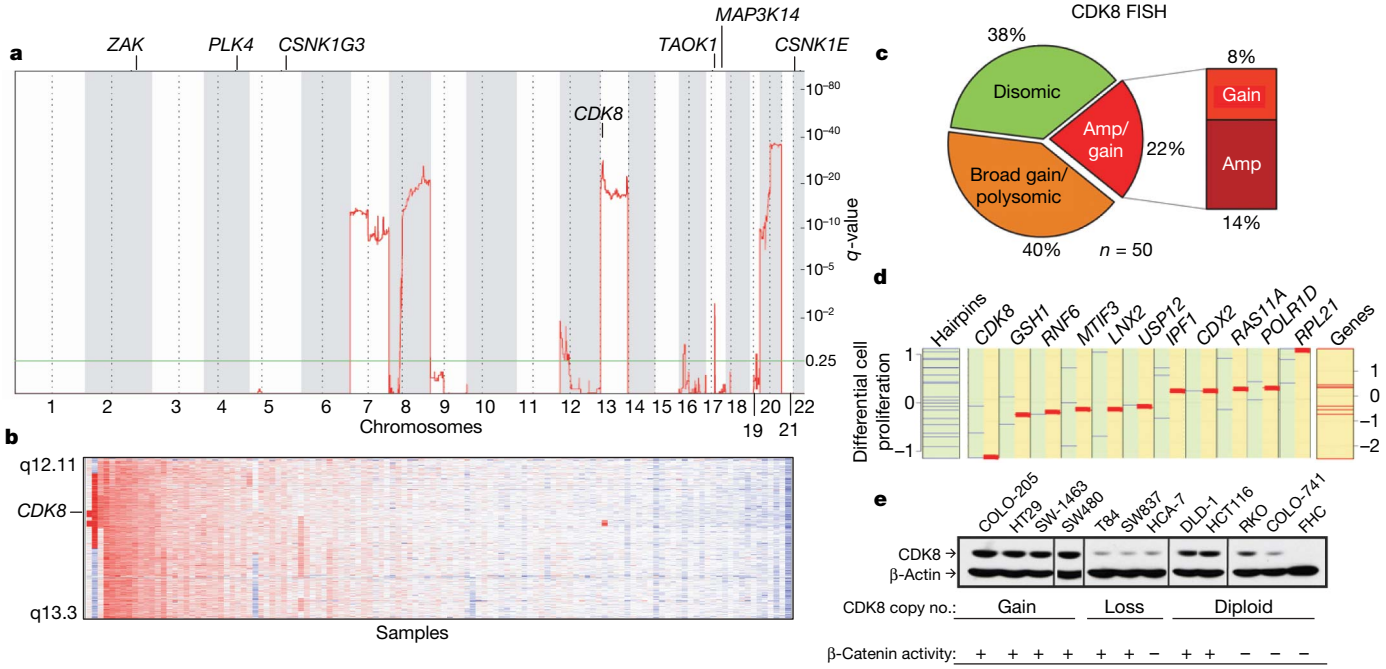
[1]Department of Medical Oncology, Dana-Farber Cancer Institute, 44 Binney Street, Boston, Massachusetts 02115, USA. [2]Department of Pathology, [3]Department of Medicine and [4]Department of Neurosurgery, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts 02115, USA. [5]Department of Neuro-oncology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. [6]Center for Cancer Genome Discovery, Dana-Farber Cancer Institute, 44 Binney Street, Boston, Massachusetts 02115, USA. [7]Broad Institute of Harvard and M.I.T., 7 Cambridge Center, Cambridge, Massachusetts 02142, USA. [8]Department of Medical Oncology, Hospital Vall d'Hebron, Passeig Vall d'Hebron, 119-129, 08035 Barcelona, Spain.

and depend on β-catenin for proliferation[6], to stably express 'pTOPFLASH' β-catenin–luciferase and 'pFOPFLASH' mutant-*Renilla* reporter constructs[7,8] (DLD-1^{Rep}). Suppression of β-catenin expression in DLD-1^{Rep} cells by three β-catenin-specific short hairpin RNAs (shRNA) markedly reduced the TOPFLASH/FOPFLASH ratio (Fig. 1a), confirming that reporter activity requires β-catenin expression. We then screened DLD-1^{Rep} cells with a shRNA library containing 4,849 shRNAs that target 1,000 genes, including 95% of all human kinases[7]. We found 34 genes which had expression that was necessary for β-catenin activity, including two known β-catenin regulators, *CSNK1G3* (ref. 9) and *CSNK1E* (ref. 10; Fig. 1b and Supplementary Table 1).

In parallel, we performed an arrayed, kinase-enriched shRNA screen in another β-catenin-dependent colon cancer cell line, HCT116, to identify genes essential for cancer cell proliferation. We identified 166 candidate genes necessary for proliferation using the criteria that at least two shRNAs targeting the same gene induced a decrease in proliferation. Among the genes identified in this screen were the oncogenes *KRAS* and *MYC* (Fig. 1c and Supplementary Table 2). Combining the genes from the two screens showed that there were nine genes for which suppression affected both β-catenin transcriptional activity and colon cancer cell proliferation (Fig. 1d).

To determine whether any of these genes are amplified in colon cancers, we used single nucleotide polymorphism (SNP) arrays and the GISTIC[11,12] statistical method to conduct a genome-wide analysis of autosomal copy number alterations in primary resection specimens from 123 human colorectal adenocarcinomas (Fig. 2a). Among the nine genes identified by our RNAi screens, only *CDK8* resides in a significant amplicon at 13q12.13–13q12.2 (false discovery rate $= 1 \times 10^{-29}$; Fig. 2a). These findings confirm recent reports that a large portion of chromosome 13 is amplified in colon cancers[13,14]. A total of 58 out of 123 (47%) samples harboured this region of copy number gain (Fig. 2b and Supplementary Table 3).

To confirm these findings, we performed fluorescence *in situ* hybridization (FISH), using probes specific for *CDK8* and *RB1* (chromosome 13 control probe), on a tissue microarray (TMA) carrying



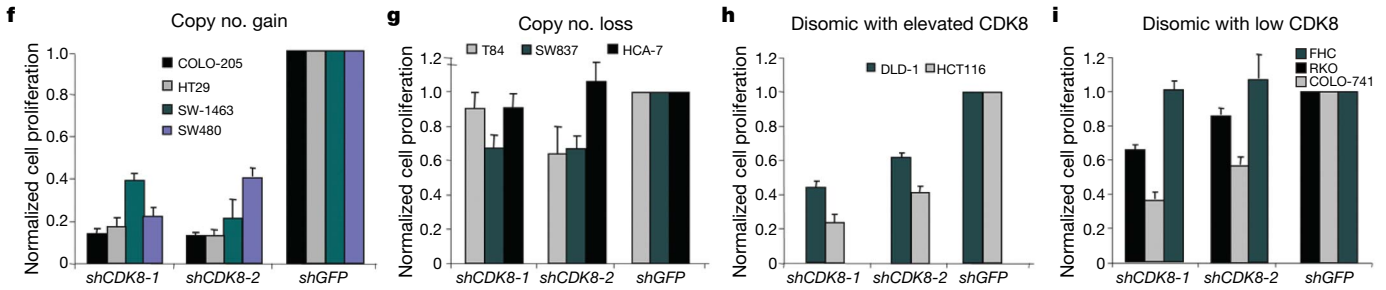


Figure 2 | Amplification and overexpression of CDK8 defines a subset of colon cancers. a, The significance (*y* axis; *q*-value) of recurrent amplifications at loci across the 22 autosomes (*x* axis) identified by GISTIC analysis. The chromosomal location of RNAi hits is indicated above the plot. **b**, Heat map showing clustering of SNP array data, on the basis of chromosome 13q12 copy number in 123 colon cancer specimens. Red indicates allelic gain, blue denotes allelic loss. **c**, Pie chart depicting the percentage of colon tumours with *CDK8* copy number gain, chromosome 13 polysomy or disomy ($n = 50$) by FISH. Cutoff criteria for FISH are shown in Supplementary Table 4. Amp, genomic amplification. **d**, GSEA comparative analysis of suppressing resident genes in the minimal 13q12 region. Blue lines represent differential scores of cell proliferation effects for each validated shRNA that suppressed target gene expression by more than 70%. The left-most column demonstrates the total pool of validated hairpins for all genes on 13q12. The red lines represent a normalized enrichment score for each gene that takes into account the cell proliferation effects of all shRNAs and scores the specificity of the effects in cell lines that harbour or lack 13q gain. **e**, Immunoblot analysis of CDK8 expression in 12 colon cancer cell lines; β-actin was used as a loading control. **f–i**, Effect of *CDK8* suppression on proliferation of colon cancer cells that harbour chromosome 13 copy number gain (**f**), loss (**g**), are disomic with increased CDK8 protein expression (**h**) and are disomic with reduced CDK8 protein expression (**i**). Bar graph depicts cell proliferation normalized to the *shGFP* control for triplicate determinations. Error bars represent mean ± s.d. for a representative experiment performed in triplicate.

50 evaluable colon cancer specimens. We detected *CDK8* copy number gain in 31 out of 50 (62%) cases. Of these 50 cases, 20 showed gains in both *CDK8* and *RB1*, indicative of polysomy and consistent with recent observations linking RB, E2F1, β-catenin and CDK8 (ref. 15). We also found *CDK8* amplifications in seven of these tumours (defined as a CDK8:control ratio ≥2) and low to moderate level copy number gain (a CDK8:control ratio >1 and <2) in a further four specimens (Fig. 2c and Supplementary Table 4). Immunohistochemical analysis of CDK8 expression in the same 50 specimens showed elevated protein levels in 13 out of 50 (26%) colon cancer samples, including those that showed *CDK8* copy number gain (Supplementary Fig. 1 and Supplementary Table 4). These observations indicate that *CDK8* is amplified and overexpressed in a substantial fraction of colon cancers.

The minimal region shared by these tumours encompasses 16 annotated genes (Supplementary Fig. 2a). To determine whether *CDK8* is the primary target of this amplicon, we first assessed expression of these genes in colon cancer cells harbouring chromosome 13q copy number gain and found that four of the genes were not expressed (Supplementary Fig. 2b). We suppressed the expression of the remaining 12 genes in four cell lines: two (HT29 and COLO-205) that harbour 13q copy number gain, and two (SW837 and T84) that show 13q loss (Supplementary Fig. 3a, b). To analyse the screen results on a per-gene basis in cell lines with either 13q12 copy number gain or deletion, we used an adaptation of the GSEA[16] method and found that *CDK8* was the only gene in this region that was required for proliferation of cell lines harbouring 13q gain (false discovery rate = 0.24; Fig. 2d and Supplementary Table 5). These observations suggested that colon cancer cells harbouring 13q12.2 amplification are particularly dependent on CDK8 expression for proliferation.

We then analysed *CDK8* copy number gain and protein expression in a panel of 12 colon cancer lines. Four (COLO-205, HT29, SW1463 and SW480)[17] of these twelve lines were found to harbour *CDK8* gain (Supplementary Fig. 3a, b), and these cell lines showed the highest amounts of CDK8 protein (Fig. 2e). Two additional colon cancer cell lines disomic for *CDK8* (DLD-1 and HCT116) also had elevated CDK8 protein amounts (Fig. 2d and Supplementary Fig. 3a, b). Suppression of *CDK8* expression induced substantially decreased proliferation in all six cell lines with elevated CDK8 protein concentrations (Fig. 2f, g) but inhibited proliferation rates in the cell lines with lower CDK8 protein levels to a lesser degree (Fig. 2h, i). Suppressing CDK8 in colon cancer cells reduced the fraction of cells in G1 and S phase, increased the number of aneuploid cells and markedly slowed cell proliferation without inducing apoptosis (Supplementary Fig. 4), similar to what was observed on suppression of β-catenin. These observations demonstrate that colon cancer cells that express elevated CDK8 levels are highly dependent on its expression for proliferation.

To determine whether CDK8 induces cell transformation, we overexpressed wild-type CDK8 or a previously reported kinase-inactive substitution mutant (D173A; CDK8-KD)[18] in immortal murine fibroblasts (NIH 3T3 cells; Fig. 3a). CDK8 expression induced focus formation, anchorage-independent colony growth and tumour formation in immunodeficient animals (Fig. 3b–e), whereas the CDK8-KD mutant failed to transform the cells. These observations confirm that CDK8 is a bona fide oncogene, the kinase activity of which is necessary for oncogenic activity.

To dissect the relationship between CDK8 and β-catenin activity, we measured endogenous β-catenin activity in the 12 cell lines used above. The RKO, COLO-741, HCA-7 and FHC cell lines do not harbour known APC or β-catenin mutations[19–21] and, as predicted, had low amounts of β-catenin activity. Of these four cell lines, suppression of *CDK8* induced a substantial proliferation effect only in COLO-741 (Fig. 2h and Supplementary Fig. 5a). Similarly, of the 12 cell lines tested, the six cell lines with the highest CDK8 elevation showed a greater dependence on β-catenin for proliferation (Supplementary Fig. 5b).

CDK8 is a cyclin-dependent kinase member of the mediator complex, which couples transcriptional regulators to the basal transcriptional machinery[4]. To explore the role of CDK8 in modulating β-catenin transcriptional activity, we confirmed that suppressing CDK8 with two independent, *CDK8*-specific shRNAs (*shCDK8-1* and *shCDK8-2*) in another cell line, SW480, also reduced β-catenin-dependent transcriptional activity (Fig. 4a). CDK8 kinase activity depends on the co-factor cyclin C[22], and we found that cyclin C knockdown preferentially affected colon cancer cell lines with chromosome 13q gain (Supplementary Fig. 6a, b). To test whether CDK8 kinase activity is required to regulate β-catenin activity, we expressed wild-type CDK8 or CDK8-KD in DLD-1^{Rep} cells carrying a shRNA targeting the 3′-untranslated region, *shCDK8-1*, and found that only wild-type CDK8 partially rescued the effects of suppressing endogenous *CDK8* (Supplementary Fig. 6c). These observations demonstrate that the kinase activity of CDK8 is necessary for both CDK8-induced transformation and β-catenin driven transcription.

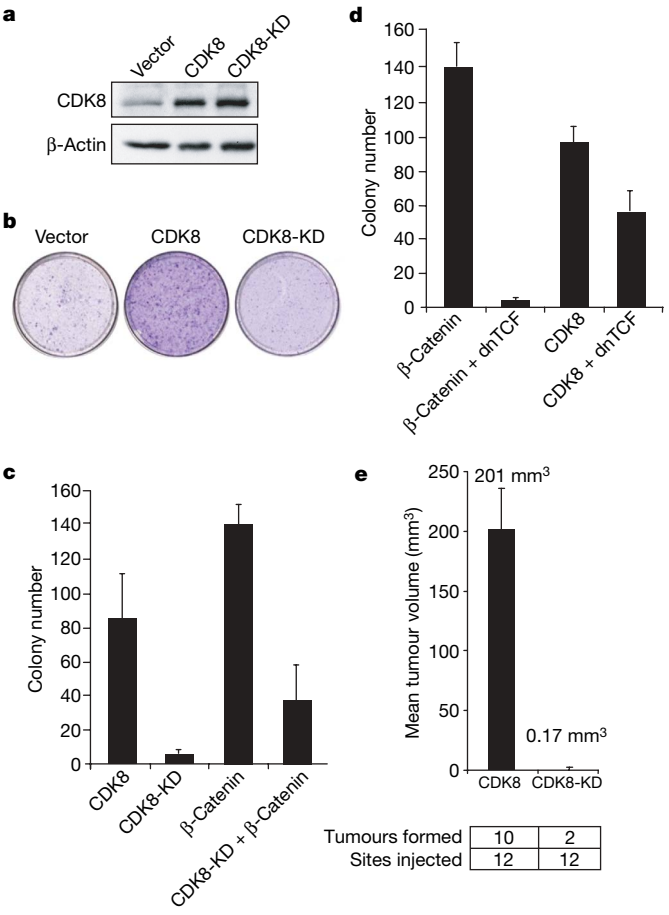


Figure 3 | CDK8 and transformation. **a**, Immunoblot analysis showing CDK8 and CDK8 kinase-dead (CDK8-KD) expression in NIH 3T3 cells. **b**, Focus formation assay of NIH 3T3 cells expressing CDK8 or CDK8-KD. **c**, CDK8 kinase activity drives anchorage-independent growth and is necessary for β-catenin (CTNNB1) mediated anchorage-independent growth. Anchorage-independent colony growth in NIH 3T3 cells infected with the indicated retroviral vectors is shown. **d**, β-Catenin suppression only partially blocks CDK8-induced anchorage-independent growth. Dominant negative TCF (dnTCF) was introduced in the presence of β-catenin or CDK8. Anchorage-independent colony growth was performed as indicated in **c**. **e**, CDK8 kinase activity drives tumour formation. Mean tumour volume from subcutaneous tumours formed NIH 3T3 cells expressing CDK8 or CDK8-KD constructs in immunodeficient mice. The difference in tumour formation between CDK8 and CDK8-KD was statistically significant, as assessed by an unpaired *t*-test (*P* value = 0.0001). All experiments were performed in triplicate and mean ± s.d. are shown.

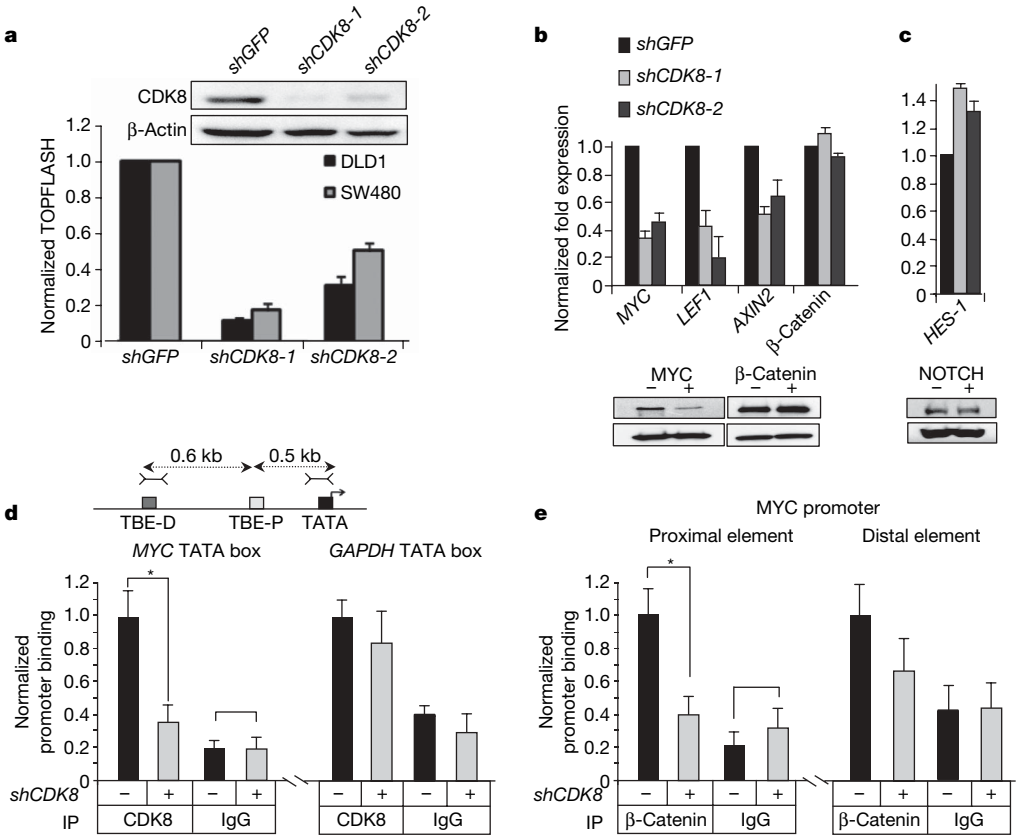


Figure 4 | CDK8 mediates transcription of β-catenin-driven downstream target genes. a, Bar graph depicts β-catenin activity as normalized TOPFLASH (TOPFLASH/FOPFLASH) after introduction of *CDK8*-specific shRNAs. An immunoblot shows CDK8 protein amounts at the time of the assay; β-actin was used as a loading control. **b**, **c**, Bar graphs show messenger RNA abundance of endogenous β-catenin and NOTCH transcriptional targets after CDK8 suppression in COLO-205 and DLD-1 cells. MYC, β-catenin and NOTCH protein expression changes were assessed by immunoblotting; β-actin expression was used as a loading control. **d**, CDK8 binds the *MYC* promoter. Schematic representation depicts the location of proximal and distal TCF binding elements (TBE, grey boxes) and *MYC* promoter TATA box. Dashed lines indicate the distance from TBEs to the TATA box. Hatched lines depict the PCR primers used for ChIP. The bar graph shows CDK8 binding to the *MYC* TATA box in COLO-205 colon cancer cells treated with the indicated shRNAs and assayed by ChIP. **e**, Bar graph depicts β-catenin binding to the indicated *MYC* promoter elements in the presence or absence of CDK8. Asterisks indicate statistical significance as assessed by an unpaired Student's *t*-test ($P \leq 0.013$ for CDK8 and $P \leq 0.14$ for the IgG control in **d**; $P \leq 0.005$ for CDK8 and $P \leq 0.34$ for the IgG control in **e**). ChIP assays were performed as described earlier. All experiments were conducted at least twice in triplicate. Error bars represent the mean ± s.d. for a representative experiment performed in triplicate. IP, immunoprecipitate.

The TCF–β-catenin complex regulates expression of several genes implicated in colon cancer, including *MYC*[23], *AXIN2* (ref. 24) and *LEF1* (ref. 25). Suppression of CDK8 in DLD-1 and COLO-205 cells reduced expression of each of these genes (Fig. 4b). In contrast, we failed to observe changes in the expression of NOTCH or HES-1 (Fig. 4c), previously reported targets of CDK8 (ref. 26). Thus, CDK8 modulates a subset of β-catenin-driven genes previously implicated in cancer[23,24,27].

We then performed chromatin immunoprecipitation (ChIP) near two verified β-catenin/TCF binding elements (TBE)[25] in the *MYC* promoter, as an example of a β-catenin regulated gene, to test whether CDK8 directly modulates *MYC* expression at the promoter level. We found that CDK8 associated with the *MYC* promoter (Fig. 4d). We therefore asked if loss of CDK8 binding at the *MYC* promoter affects the ability of β-catenin to bind at the proximal and distal TBEs. Suppression of *CDK8* expression reduced the amount of β-catenin bound to the proximal element in the *MYC* promoter but had little effect on the amount associated with the distal element (Fig. 4e). These observations implicate CDK8 and the mediator complex[4] as a direct regulator of β-catenin-driven transcription.

To test whether CDK8 activity is also required for β-catenin-driven transformation, we expressed the dominantly interfering CDK8-KD mutant[28] in transformed NIH 3T3 cells expressing a constitutively active β-catenin mutant (Fig. 3c). Disruption of CDK8 activity inhibited β-catenin-driven transformation, whereas a dominantly interfering TCF construct, previously shown to inhibit β-catenin-induced cellular transformation[29], only partially abrogated CDK8-mediated transformation (Fig. 3d). These observations suggest that although CDK8 is required for β-catenin-mediated transformation, the full capacity of CDK8 to transform cells may extend beyond its ability to activate β-catenin.

We have used an integrated approach to identify *CDK8* as an oncogene in a substantial fraction of colorectal cancers and demonstrate that the kinase activity of CDK8 is essential for it to be able to regulate β-catenin-dependent transcription and transformation. These observations indicate that CDK8 acts in part by co-activating β-catenin-driven transcription in colon cancers characterized by both high CDK8 expression and β-catenin activity. Accordingly, therapeutic interventions that target the CDK8 kinase activity in such cancers may be of clinical value.

METHODS SUMMARY

Lentiviral infections were performed using pLKO.1 lentiviral shRNA constructs generated by the RNAi consortium[7] and are listed in Supplementary Table 6. For high-throughput screening, cells infected with lentiviruses were allowed to grow for either 4 days (DLD-1^{Rep}) or 5 days (HCT116). Suppression of β-catenin activity was defined as the ability of at least one shRNA to decrease activity more than 2 s.d. below the *Z* score. Suppression of proliferation was defined as the capacity for at least one shRNA to decrease proliferation more than 1.5 s.d. below the mean *Z* score and at least one extra shRNA targeting the same kinase to decrease proliferation more that 1 s.d. below the mean *Z* score. For tissue

analyses, a TMA composed of human colon cancer tissue and matching patient normal colon was subjected to immunohistochemical staining. For FISH, BAC clones were hybridized in dual colours to 4-μm TMA sections as follows: the RB1 probe labelled in SpectrumGreen was used as a chromosome 13 control probe and the RP11-726I20 BAC, spanning CDK8, was labelled with SpectrumOrange dUTP (both from Abbott Molecular/Vysis).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 30 March; accepted 19 June 2008.
Published online 14 September 2008.

1. Bienz, M. & Clevers, H. Linking colorectal cancer to Wnt signaling. *Cell* **103,** 311–320 (2000).
2. Camp, R. L., Chung, G. G. & Rimm, D. L. Automated subcellular localization and quantification of protein expression in tissue microarrays. *Nature Med.* **8,** 1323–1327 (2002).
3. Vogelstein, B. *et al.* Genetic alterations during colorectal-tumor development. *N. Engl. J. Med.* **319,** 525–532 (1988).
4. Conaway, R. C., Sato, S., Tomomori-Sato, C., Yao, T. & Conaway, J. W. The mammalian Mediator complex and its role in transcriptional regulation. *Trends Biochem. Sci.* **30,** 250–255 (2005).
5. Smith, G. *et al.* Mutations in APC, Kirsten-ras, and p53—alternative genetic pathways to colorectal cancer. *Proc. Natl Acad. Sci. USA* **99,** 9433–9438 (2002).
6. van de Wetering, M. *et al.* The β-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell* **111,** 241–250 (2002).
7. Moffat, J. *et al.* A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* **124,** 1283–1298 (2006).
8. Korinek, V. *et al.* Constitutive transcriptional activation by a β-catenin–Tcf complex in $APC^{-/-}$ colon carcinoma. *Science* **275,** 1784–1787 (1997).
9. Davidson, G. *et al.* Casein kinase 1γ couples Wnt receptor activation to cytoplasmic signal transduction. *Nature* **438,** 867–872 (2005).
10. Hino, S., Michiue, T., Asashima, M. & Kikuchi, A. Casein kinase Iε enhances the binding of Dvl-1 to Frat-1 and is essential for Wnt-3a-induced accumulation of β-catenin. *J. Biol. Chem.* **278,** 14066–14073 (2003).
11. Beroukhim, R. *et al.* Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. *Proc. Natl Acad. Sci. USA* **104,** 20007–20012 (2007).
12. Weir, B. A. *et al.* Characterizing the cancer genome in lung adenocarcinoma. *Nature* **450,** 893–898 (2007).
13. Martin, E. S. *et al.* Common and distinct genomic events in sporadic colorectal cancer and diverse cancer types. *Cancer Res.* **67,** 10736–10743 (2007).
14. Tsafrir, D. *et al.* Relationship of gene expression and chromosomal abnormalities in colorectal cancer. *Cancer Res.* **66,** 2129–2137 (2006).
15. Morris E. J. . *et al.* E2F1 represses β-catenin transcription and is antagonized by both pRB and CDK8. *Nature* doi:10.1038/nature07310 (this issue).
16. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102,** 15545–15550 (2005).
17. Garraway, L. A. *et al.* Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. *Nature* **436,** 117–122 (2005).
18. Gold, M. O. & Rice, A. P. Targeting of CDK8 to a promoter-proximal RNA element demonstrates catalysis-dependent activation of gene expression. *Nucleic Acids Res.* **26,** 3784–3788 (1998).
19. Ilyas, M., Tomlinson, I. P., Rowan, A., Pignatelli, M. & Bodmer, W. F. β-Catenin mutations in cell lines established from human colorectal cancers. *Proc. Natl Acad. Sci. USA* **94,** 10330–10334 (1997).
20. Sparks, A. B., Morin, P. J., Vogelstein, B. & Kinzler, K. W. Mutational analysis of the APC/β-catenin/Tcf pathway in colorectal cancer. *Cancer Res.* **58,** 1130–1134 (1998).
21. Rowan, A. J. *et al.* APC mutations in sporadic colorectal tumors: A mutational ''hotspot'' and interdependence of the ''two hits''. *Proc. Natl Acad. Sci. USA* **97,** 3352–3357 (2000).
22. Tassan, J. P., Jaquenoud, M., Leopold, P., Schultz, S. J. & Nigg, E. A. Identification of human cyclin-dependent kinase 8, a putative protein kinase partner for cyclin C. *Proc. Natl Acad. Sci. USA* **92,** 8871–8875 (1995).
23. Sansom, O. J. *et al. Myc* deletion rescues Apc deficiency in the small intestine. *Nature* **446,** 676–679 (2007).
24. Murakami, T. *et al.* Constitutive activation of Wnt/β-catenin signaling pathway in migration-active melanoma cells: role of LEF-1 in melanoma with increased metastatic potential. *Biochem. Biophys. Res. Commun.* **288,** 8–15 (2001).
25. He, T. C. *et al.* Identification of c-*MYC* as a target of the APC pathway. *Science* **281,** 1509–1512 (1998).
26. Fryer, C. J., White, J. B. & Jones, K. A. Mastermind recruits CycC:CDK8 to phosphorylate the Notch ICD and coordinate activation with turnover. *Mol. Cell* **16,** 509–520 (2004).
27. Yook, J. I. *et al.* A Wnt–Axin2–GSK3β cascade regulates Snail1 activity in breast cancer cells. *Nature Cell Biol.* **8,** 1398–1406 (2006).
28. Chiang, M. Y. *et al.* Identification of a conserved negative regulatory sequence that influences the leukemogenic activity of NOTCH1. *Mol. Cell. Biol.* **26,** 6261–6271 (2006).
29. Kolligs, F. T., Hu, G., Dang, C. V. & Fearon, E. R. Neoplastic transformation of RK3E by mutant β-catenin requires deregulation of Tcf/Lef transcription but not activation of c-*myc* expression. *Mol. Cell. Biol.* **19,** 5696–5706 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank E. Shin for technical assistance with immunohistochemistry, M. Miri for assistance with sample collection and G. Getz for assistance with SNP array analysis. This work was supported in part by a grant from the US NCI/NIH R33CA128625 (W.C.H.), T32 NIH grant (R.F.) and a GI SPORE Career Development Grant (P50CA127003; R.F.), a Harvard-MIT Clinical Investigator Training Program Fellowship (A.J.B), Department of Defense Prostate Cancer Postdoctoral Fellowships (I.G. and S.Y.K.), Warren-Whitman-Richardson, Hagerty Foundation Research Fellowships (I.F.D.) and K12 award (M.G.C.). J.B. is a Beatriu de Pinos Fellow of the Departament d'Educació i Universitats de la Generalitat de Catalunya.

Author Contributions R.F., A.J.B., M.M. and W.C.H. designed the approach. R.F., A.J.B., E.B., E.F., I.F.D., I.G., J.S.B., N.V., S.J.S., S.O., S.Y.K. and Y.S. performed the experiments. J. Barretina, J. Baselga, J.T. and R.A.S. provided samples and analysis. A.H.L., A.J.B., C.M., C.S.F., D.E.R., E.B., E.F., I.F.D., I.G., J.A.C., M.G.C., M.L., M.M., P.T., R.A.S., R.F., S.F., S.J.S., S.O., S.Y.K. and W.C.H. analysed data. R.F., A.J.B., R.A.S. and W.C.H. wrote the manuscript.

Author Information The SNP data can be found at http://research3.dfci.harvard.edu/cdk8colon/index.php. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Correspondence and requests for materials should be addressed to W.C.H. (william_hahn@dfci.harvard.edu).

METHODS

Plasmids and antibodies. Human CDK8 and CDK8 (D173A) kinase-dead mutants were obtained in a pCMV2 vector backbone (gift from A. Rice) and cloned into a pBABE-puro-Flag retroviral vector. pBABE-puro-S33Y-β-catenin has been previously described[1]. pTOPFLASH and pFOPFLASH were generated by inserting a tandem 8× TOPFLASH or 8× FOPLFASH element into vectors pGL4.15 (Promega) and pGL4.80 (Promega), respectively. The following antibodies were used for all experiments: CDK8 (goat anti-hCDK8, Santa Cruz, sc-1521); MYC (rabbit anti-c-Myc, Santa Cruz, sc-788); p21 (rabbit anti-p21, Santa Cruz, sc-397); Notch (rabbit anti-Notch1, Epitomics, 1935); β-catenin (mouse anti-β-catenin, BD Transduction, 610154); β-actin (HRP conjugated β-actin; Santa Cruz, sc47778); goat IgG (Santa Cruz, sc-2028).

Cell culture and gene transfer. DLD-1, COLO-741, RKO, HCT116, SW-1463, T84, COLO-205, SW480, HT29, SW837 and HCA-7 cell lines were maintained in DMEM supplemented with 10% fetal bovine serum (Sigma). FHC cells were grown in DMEM and F12 medium (ATCC) supplemented with 10% fetal bovine serum (Sigma). All transfections were performed using Fugene 6 (Roche) and retroviral and lentiviral infections were performed as described[30]. To generate stable cell lines, cells were selected using 2 μg ml^{-1} of puromycin for 3 days, 200 μg ml^{-1} of hygromycin for 5 days or 500 μg ml^{-1} of G418 for 7 days. Clonal DLD-1^{Rep} and SW480^{Rep} cells were generated by serial transfection of DLD-1 and SW480 cells with the pTOPFLASH plasmid and pFOPFLASH plasmids, selection, and isolation of single cell clones after 14 days of growth in media with the appropriate resistance markers.

Focus formation, anchorage-independent growth and tumorigenicity assays. Anchorage-independent growth and tumorigenicity assays were performed as described[31]. Colonies from soft agar assays were assessed using images taken from plates examined under a dissection microscope and analysed using ImageJ software.

Expression analysis and chromatin immunoprecipitation. Immunoblotting was performed according to standard methods. For RT–PCR, total RNA was isolated using RNeasy (Qiagen) and cDNA synthesis was performed using the Advantage RT/PCR kit (Clontech). For quantitative RT–PCR, we used SYBR Green (Applied Biosystems). ChIP experiments were performed using EZ-CHIP (Upstate Biotech).

RNAi and reporter assays. Lentiviral infections of cell lines were performed using pLKO.1 lentiviral shRNA constructs generated by the RNAi consortium, which are listed in Supplementary Table 6. A pLKO.1 *shGFP* control plasmid containing the following targeting sequence was used: GCCCGCAAGCTGAC CCTGAAGTTCATTCAAGAGATGAACTTCAGGGTCAGCTTGCTTTTT. For high-throughput screening, cells infected with lentiviruses were allowed to grow for 4 days (DLD-1^{Rep}) or 5 days (HCT116). Dual Glo Luciferase (Promega) and Cell Titer Glo (Promega) were used to measure luciferase activity for DLD-1^{Rep} β-catenin activity and HCT116 proliferation, respectively. *Z* scores were calculated as described[32]. Suppression of β-catenin activity was defined as the ability of at least one shRNA to decrease activity more than 2 s.d. below the *Z* score. Suppression of proliferation was defined as the capacity for at least one shRNA to decrease proliferation more than 1.5 s.d. below the mean *Z* score and at least one extra shRNA targeting the same kinase to decrease proliferation more that 1 s.d. below the mean *Z* score. RNAi analysis of resident 13q12 genes essential for colon cancer proliferation was conducted using colon cancer cell lines that harbour (HT29 and COLO-205) or lack (SW837 and T84) chromosome 13 copy number gain. These cells were infected in parallel with lentiviruses harbouring at least three different shRNA targeting each gene and subsequent cell proliferation and gene expression was measured. shGSEA analysis was then conducted to derive a statistical score of proliferative effects.

Tumour collection, processing and SNP array analysis. DNA was collected from freshly frozen colorectal adenocarcinomas derived from primary resection specimens (123) and matched normal colonic tissues (115). Whenever possible, samples with estimated tumour content of more than 70% on the basis of pathological review were chosen for this analysis. Samples were obtained from the following five sites: the Massachusetts General Hospital (25), the Brigham and Women's Hospital/Dana Farber Cancer Institute (46), the Vall d'Hebron Hospital in Barcelona (9) and the Genomics Collaborative Incorporated (43). From each sample, 250 ng of unamplified DNA was digested with StyI (New England Biolabs), ligated to a universal adaptor using T4 ligase (New England Biolabs) and PCR amplified using Titanium Taq (Clontech). Using Affymetrix protocols, amplified DNA underwent clean-up steps and was then fragmented with DNaseI (Affymetrix), labelled, denatured and hybridized to the StyI chip of the Affymetrix 500K Human Mapping Array set. The GeneChip Scanner 3000 7G (Affymetrix) was used to scan all hybridized arrays. Samples were run in 96-sample batches which included 6 standard control DNA input samples as well as DNA from tumours and normal tissues of different histologies. Raw array data from samples in each batch were processed using the GenePattern software package (http://www.broad.mit.edu/cancer/software/genepattern/) to create copy number values for each SNP. All mapping information for each SNP, cytoband and gene was on the basis of Affymetrix annotations and the hg17 (May 2004) genome build (http://genome.ucsc.edu). See Supplementary Methods for further details.

Immunohistochemical analysis and fluorescence *in-situ* hybridization (FISH). A TMA composed of 88 cases of human colon cancer tissue and matching patient normal colon was obtained from the Center for Molecular Oncologic Pathology (Brigham and Women's Hospital/Dana Farber Cancer Institute). Thirty-eight cases were excluded on the basis of inadequate or insufficient tissue for comparative analysis with FISH. The TMAs were subjected to citrate buffer and pressure cooker antigen retrieval and immunostained with the CDK8 antibody (1:100 dilution). Tissues were assessed for morphology and expression by a pathologist (S.F.). Four-micrometre TMA sections were mounted on standard glass slides for FISH. BAC clones were hybridized in dual colours as follows: the RB1 probe (Abbott Molecular/Vysis) was used as a chromosome 13 control probe and was obtained labelled in SpectrumGreen, and the RP11-726I20 BAC, spanning *CDK8*, was labelled with SpectrumOrange dUTP (Abbott Molecular/Vysis).

Statistical analysis of genomic copy number alterations and GISTIC. In brief, raw SNP data are renormalized by comparison of the SNP signal intensity in the tumour to that of the average of the five closest normal samples. Renormalized data are then smoothed and segmented and corrected for batch effects and published copy number polymorphisms. To identify recurrent amplifications above a threshold $\log_2$ copy number of 0.1, each locus is first given a score on the basis of frequency and intensity of amplification. To determine the false discovery rate (*q*-value), the score at each SNP is compared against that achieved by randomly permuting the copy number data. Detailed methods of the GISTIC algorithm are found elsewhere[31,33].

Derivation of GSEA for statistical analysis of shRNA effects. To analyse the screen results on a per-gene basis and compare the effects of genes between cell lines with 13q12 copy number gain and deletion, we adapted the GSEA method to evaluate enrichment of gene sets in gene expression analysis[14]. In analogy with GSEA, the 'gene set' is the set of shRNA hairpins that represent a gene. We first standardize each sample shRNA construct value using the median and maximum absolute deviation of a collection of shRNA control hairpins expressed in the same cell line. The hairpins were then mapped across samples and sorted according to the magnitude of their differential cell proliferation score (that is, the difference of their means in each phenotype: 13q12 copy number gain versus control). Only shRNAs that suppressed their target gene by more than 70% were analysed for cell proliferation effects. Once the effective hairpins have been sorted this way, an 'enrichment' score is computed for each gene on the basis of the distribution of the shRNAs in the list. The enrichment score is computed using a two-sample statistic on the basis of the likelihood ratio and is representative of both the extremeness of the hairpins' differential cell proliferation scores of a given gene, and also of their consistency. The lower the hairpins of a given gene appear in the list, the higher their (negative) enrichment score and the more the decrease in proliferation the gene produces between the samples with and without 13q12 copy number gain. To account for the fact that different genes have different number of hairpins, we normalized each enrichment score using random permutations of a hairpin set of the same size. This permutation test also provides nominal *P* values for each gene-enrichment score. The end result is a list of genes sorted by their normalized enrichment scores and a set of complementary estimates of statistical significance such as nominal, family wise and Bonferroni *P* values plus a false discovery rate.

Fluorescence *in situ* hybridization. Four-micrometre TMA sections were mounted on standard glass slides and baked at 60 °C for at least 2 h. Slides were soaked for 15 min in xylene that was pre-warmed to 55 °C and then in room temperature xylene for a further 15 min. Slides were then soaked in two successive 100% ethanol washes for at least 2 min each and allowed to air dry. The slides were then boiled in 100 mM Tris, 50 mM EDTA, pH 7.0, for 1 h, and soaked in 2× standard sodium citrate (SSC) for 5 min. Slides were placed on a 37 °C Thermobrite (StatSpin) and digested with Digest-All III solution (Invitrogen) for two 15 min digestions. Samples were fixed for 2 min in 10% phosphate buffered formalin, followed by dehydration for 2 min each in 70%, 80%, 90% and 100% ethanol.

One microgram of BAC DNA was labelled using a nick-translation kit (Abbott Molecular/Vysis) following the manufacturer's instructions. Labelled DNA was precipitated at −80 °C for 1 h with 10 μl Cot-1 DNA (1 mg ml^{-1} stock), pelleted, air dried and resuspended in the appropriate volume of 50% Hybrisol (50% formamide, 2× SSC, 10% dextran sulphate). Resuspended pellets were incubated at 37 °C for 1 h. The final probe concentration was approximately 50–100 ng μl^{-1}.

Slides were denatured in Coplin jars with 70% formamide/2× SSC at 94 °C for 5 min, followed by successive 2 min dehydrations each in 70% and 80% ethanol, at 4 °C, and 90% and 100% ethanol, at room temperature. Probes were denatured at 94 °C for 10 min, followed by snap cooling on ice. Denatured probes were applied to air-dried slides; coverslips were applied and then sealed with rubber cement. Hybridizations were performed for at least 16 h at 37 °C in a dark humid chamber. Slides were washed in 2× SSC at 70 °C for 10 min, rinsed at room temperature in 2× SSC and counterstained with DAPI (4′,6-diamidino-2-phenylindole, Abbott Molecular/Vysis). Slides were imaged using an Olympus BX51 fluorescence microscope. Individual images were captured using an Applied Imaging system running CytoVision Genus version 3.9.

30. Boehm, J. S. *et al.* Integrative genomic approaches identify IKBKE as a breast cancer oncogene. *Cell* **129,** 1065–1079 (2007).
31. Beroukhim, R. *et al.* Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. *Proc. Natl Acad. Sci. USA* **104,** 20007–20012 (2007).
32. Damalas, A. *et al.* Deregulated β-catenin induces a p53- and ARF-dependent growth arrest and cooperates with Ras in transformation. *EMBO J.* **20,** 4912–4922 (2001).
33. Weir, B. A. *et al.* Characterizing the cancer genome in lung adenocarcinoma. *Nature* **450,** 893–898 (2007).

LETTERS

E2F1 represses β-catenin transcription and is antagonized by both pRB and CDK8

Erick J. Morris[1]*, Jun-Yuan Ji[1]*, Fajun Yang[1,2], Luisa Di Stefano[1], Anabel Herr[1], Nam-Sung Moon[1], Eun-Jeong Kwon[3], Kevin M. Haigis[1,4], Anders M. Näär[1,2] & Nicholas J. Dyson[1]

The E2F1 transcription factor can promote proliferation or apoptosis when activated, and is a key downstream target of the retinoblastoma tumour suppressor protein (pRB). Here we show that E2F1 is a potent and specific inhibitor of β-catenin/T-cell factor (TCF)-dependent transcription, and that this function contributes to E2F1-induced apoptosis. E2F1 deregulation suppresses β-catenin activity in an adenomatous polyposis coli (APC)/glycogen synthase kinase-3 (GSK3)-independent manner, reducing the expression of key β-catenin targets including *c-MYC*. This interaction explains why colorectal tumours, which depend on β-catenin transcription for their abnormal proliferation, keep *RB1* intact. Remarkably, E2F1 activity is also repressed by cyclin-dependent kinase-8 (CDK8), a colorectal oncoprotein[1]. Elevated levels of CDK8 protect β-catenin/TCF-dependent transcription from inhibition by E2F1. Thus, by retaining *RB1* and amplifying *CDK8*, colorectal tumour cells select conditions that collectively suppress E2F1 and enhance the activity of β-catenin.

E2F1 is generally dispensable for cell proliferation but is selectively activated in response to specific cues, such as DNA damage, where it drives the expression of pro-apoptotic genes. Ectopic expression of *Drosophila E2F1* (*dE2f1*) in the developing wing causes apoptosis, giving a visible, dosage-sensitive phenotype that we have used to screen for *in vivo* regulators of E2F-dependent apoptosis[2]. Using this strategy, we found a novel interaction between *dE2f1* and the Wnt signalling pathway (Fig. 1). An apoptotic, gnarled wing phenotype (Fig. 1b), caused by elevated *dE2f1* in newly eclosed wing epithelial cells, was strongly suppressed by co-expression of the *Drosophila* β-catenin orthologue *armadillo* (*arm*; Fig. 1d), and partly suppressed by co-expression of *pangolin* (*pan*; Fig. 1e), which encodes *dTCF*, the transcription factor partner of *arm*. Moreover, expression of dominant-negative, amino-terminally truncated *pan* (*dTCFΔN*) phenocopied the *dE2f1* wing phenotype (Fig. 1f), and ectopic expression of *shaggy* (*sgg*; the GSK3 orthologue and negative regulator of Arm protein stability) strongly enhanced the effects of *dE2f1* (Fig. 1g, h). Using a stable and activated-mutant form of *arm* (*arm**; S44Y mutation in sgg/GSK3 phosphorylation site[3]), we tested whether *dE2f1* expression could modify an *arm*-dependent phenotype. Expression of *arm** under the *GMR* eye specific promoter causes a rough eye phenotype (Fig. 1k) that was partly suppressed by co-expression of *dE2f1/dDp* (Fig. 1l). Together, these genetic interactions show a strong functional antagonism between elevated dE2F1/dDP and Arm/β-catenin signalling *in vivo*.

Wnt/β-catenin signalling is important during development and regulates diverse aspects of cell function, including proliferation, differentiation and survival[4,5]. The genetic interactions suggested that E2F1 might inhibit Arm/β-catenin-dependent transcription. To test this, and to ask whether the interaction was conserved in human cells, we examined the effects of E2F1 on activation of a TCF-luciferase reporter (pTopFLASH) by a stable, tumour-derived form of β-catenin (S33Y[6]). In human Saos2 cells—a *p53*- and *Rb*-deficient cell line that exhibits low basal Wnt activity—expression of E2F1 strongly inhibited S33Y-β-catenin transcription (Fig. 1m). Control experiments showed that the inhibition was not an indirect consequence of E2F1-induced apoptosis or cell cycle progression (Fig. 1 and Supplementary Fig. 2). Inhibition by E2F1 was comparable to the effects of dominant-negative forms of TCF1 or TCF4, enhanced by co-expression of the E2F1-dimerization partner DP1, and abrogated by mutation of DP1- or DNA-binding domains of E2F1 (Supplementary Figs 2 and 3). Remarkably, as little as 10 ng of E2F1 expression plasmid repressed pTopFLASH activity tenfold when co-transfected with DP1 (Supplementary Fig. 2c). E2F1 activated transcription from a canonical E2F-luciferase reporter (pE2F4B) in these cells (Fig. 1m, n), indicating that it has context-dependent effects. Inhibition of β-catenin was a specific property of E2F1, as other activator E2Fs had little effect (Fig. 1n). Using E2F1/3 chimaeras[7], we mapped the inhibitory region to the Marked Box and adjacent domains of E2F1 (Supplementary Fig. 3), regions that allow selective interactions with other transcription factors and that determine the differences between the transcriptional signatures of E2F1 and E2F3 (ref. 8).

The pro-apoptotic activity of E2F1 has been linked to the *TP53* and *TP73* tumour suppressors; intriguingly, both of these also affect β-catenin-dependent transcription from the pTopFLASH reporter[9–11]. p53, like E2F1, inhibited pTopFLASH transcription, whereas p73, a p53-related gene that is a transcriptional target of E2F1, activated pTopFLASH (Fig. 1o). In *TP53*-deficient Saos2 cells, E2F1 repressed β-catenin-mediated pTopFLASH activity and dominantly suppressed the stimulatory effects of p73 (Fig. 1o). Thus, E2F1 is a potent inhibitor of β-catenin/TCF-activated transcription that acts independently of p53 and is dominant over p73.

To determine the effects of E2F1 on endogenous β-catenin-regulated genes we first examined *c-MYC*, one of the best-studied targets of β-catenin and a key mediator of the pro-proliferative effects of deregulated β-catenin during tumorigenesis[12]. Levels of *c-MYC* messenger RNA and protein rapidly decreased after E2F1 induction in Saos2-*TR-E2F1* cells (Fig. 2a, c). *PPARδ* and *CD44*, two other well-studied Wnt targets, showed similar changes (Fig. 2c). As a control, E2F1 strongly activated the well-known E2F target genes *CCNE1* (*Cyclin E*) and *TP73* (Fig. 2b). Similar effects on Wnt target genes were observed when E2F1 was expressed in colorectal cancer cells (Fig. 2f).

[1]Laboratory of Molecular Oncology, Massachusetts General Hospital Cancer Center and Harvard Medical School, 13th Street, Building 149, Charlestown, Massachusetts 02129, USA. [2]Department of Cell Biology, Harvard Medical School, Boston, Massachusetts 02115, USA. [3]Massachusetts General Hospital and The Vincent Center for Reproductive Biology, 55 Fruit Street, Their 901, Boston, Massachusetts 02114, USA. [4]Laboratory of Molecular Pathology, Massachusetts General Hospital Cancer Center and Harvard Medical School, 13th Street, Building 149, Charlestown, Massachusetts 02129, USA.
*These authors contributed equally to this work.

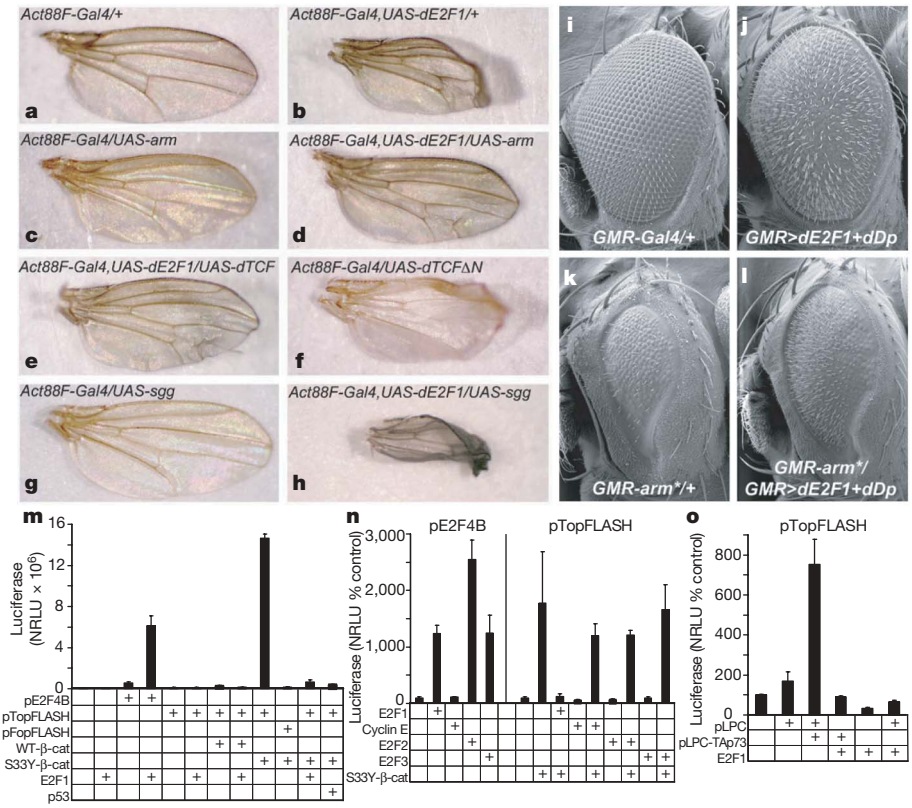


Figure 1 | Functional antagonism between E2F1 and β-catenin/TCF-signalling. a, Normal *Drosophila* wing of *Act88F-Gal4/+* genotype. A *dE2F1*-induced apoptotic wing phenotype (**b**) is strongly suppressed by co-expression of *arm* (**d**) or *pan/dTCF* (**e**), phenocopied by expression of dominant-negative *dTCF* (*dTCFΔN*) (**f**), and strongly synergizes with *sgg/GSK3* co-expression (**h**). Expression of *arm* (**c**), *sgg* (**g**) or *pan* (not shown) alone does not induce a wing phenotype. **i–l**, *GMR*-mediated eye-specific expression of *dE2F1/dDp* strongly suppresses a rough eye phenotype induced by activated-*arm* (*arm**; *S44Y*) expression. All phenotypes were compared in female progeny from F_1 crosses conducted at 25 °C versus control (w^{1118} background). **m**, E2F1 expression activates a canonical pE2F4B-luciferase reporter while abrogating the S33Y-β-catenin-mediated activation of pTopFLASH. **n**, Inhibition of pTopFLASH activity is specific to E2F1. **o**, E2F1 dominantly inhibits pTopFLASH activation by p73. All data are expressed as mean ± s.d. ($n = 3$) of normalized relative light units (NRLU) of luciferase.

The list of known Wnt targets includes genes that control β-catenin degradation, such as *AXIN1* and *AXIN2*. Interestingly, the expression of *AXIN1* and *AXIN2*, as well as *SIAH1*, a p53-inducible, GSK3-independent promoter of β-catenin degradation[13,14], were all significantly activated by E2F1 (Fig. 2d, e), consistent with previous studies suggesting that *AXIN2* and *SIAH1* are E2F-target genes[15,16]. Accordingly, the level of β-catenin protein decreased at later time points following E2F1 expression, a change that preceded apoptosis (Fig. 2g and Supplementary Fig. 4). Similarly, ectopic expression of *Drosophila* dE2F1/dDP reduced Arm protein levels (Supplementary Fig. 4h).

The mechanism of E2F1-dependent β-catenin downregulation is probably distinct from the changes observed during epithelial–mesenchymal transition, or following the disruption of adherens junctions or focal-adhesions, as markers for these processes were unperturbed by E2F1 expression (Supplementary Fig. 4i). Instead, E2F1 induced the post-translational degradation of β-catenin in a GSK3- and caspase-independent fashion (Fig. 2h and Supplementary Fig. 4g). E2F1-mediated degradation of β-catenin is functionally significant because re-expression of stable, tumour-derived mutants of β-catenin, or treatment with GSK3-inhibitors, partly abrogated E2F1-dependent apoptosis (Fig. 2i). Taken together, these results show that E2F1 inhibits β-catenin activity via transcriptional antagonism and β-catenin degradation, and that this inhibition contributes to E2F1-induced apoptosis.

β-Catenin-dependent transcription is crucially important for cell proliferation in colorectal cancer cells. Mutations in *APC* or *CTNNB1* (β-catenin) occur early in colorectal tumorigenesis, leading to pre-malignant polyps. Additional mutations contribute to the transition to malignant adenocarcinoma[4,5]. An unusual feature of colorectal cancer cells is that they rarely (if ever) acquire mutations in the *RB1* tumour suppressor gene. Paradoxically, *RB1* copy gains are frequently found in colorectal cancer cells, often resulting in protein overexpression[17]. Conditional inactivation of murine *Rb* by *Villin-Cre* leads to aggressive tumours in various tissues, but rarely in the gastrointestinal tract[18,19], and the knockdown of pRB reduces cell proliferation and anchorage-independent growth of human colon cancer cell lines[20]. Accordingly, we found elevated levels of pRB and β-catenin co-localized within the epithelium of Apc^{Min} colonic tumours in mice (Fig. 3a, b). We hypothesized that colorectal tumour cells might select for mechanisms that limit the activity of E2F1, and that in this context the pRB tumour suppressor might act to sustain high levels of β-catenin/TCF-dependent transcription.

To test this hypothesis, we used a stable line of U2OS osteosarcoma cells containing a doxycycline (Dox)-inducible short-hairpin RNA targeting *Rb* (U2OS-*shRb*)[21]. Depletion of pRB increased transcription from an E2F reporter and inhibited transcription from the pTopFLASH reporter (Fig. 3c, d). In SW480 colorectal cancer cells that contain mutant *APC* and have deregulated β-catenin[22], the expression of E2F1 sufficed to activate the E2F-luciferase reporter and to inactivate basal pTopFLASH transcription (Fig. 3e). Moreover, a short hairpin RNA (shRNA) vector that targets pRB inhibited the activity of endogenous β-catenin/TCF (Fig. 3f) and strongly inhibited cell proliferation (Fig. 3g), an effect partly rescued by co-expression of *S33Y-β-catenin/TCF*, but not *Bcl-2* (Fig. 3g). Together, these observations provide a molecular explanation why colorectal tumour cells maintain the expression of pRB.

Because E2F1 is a potent inhibitor of β-catenin, we reasoned that tumour cells might select for additional ways to limit its activity. We generated transgenic flies that allowed us to knockdown dE2F1 in a tissue-specific manner (*dE2f1*^{RNAi}; Fig. 4 and Supplementary Fig. 5). This approach reduced dE2F1 activity *in vivo* and generated phenotypes that were used to screen for factors that were rate-limiting for dE2F1-dependent proliferation *in vivo* (see Methods). We identified *dCdk8*^{c01804}, a hypomorphic mutant allele of *Drosophila Cdk8*, as a strong and specific suppressor of *dE2f1*^{RNAi} phenotypes in both the eye and wing (Fig. 4a–c and Supplementary Fig. 5). CDK8, Cyclin C, MED12 and MED13 form a sub-module of the Mediator complex, a large multi-subunit regulator of transcription[23,24]. We observed increased expression of dE2F1-regulated genes in *dCdk8*- or *dCycC*-mutant larvae (Fig. 4d). RNA interference (RNAi)-mediated depletion of dCDK8 or dCycC in *Drosophila* SL2 cells caused similar

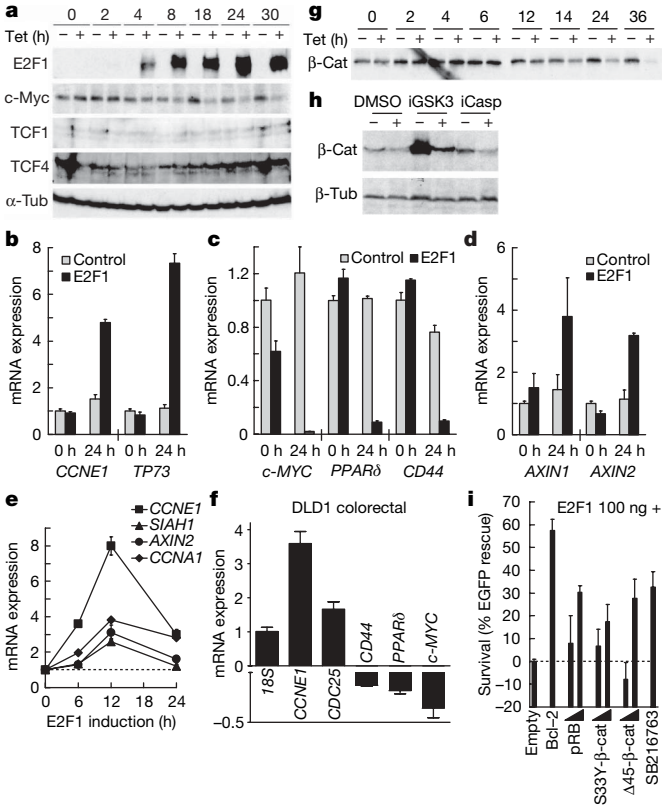


Figure 2 | E2F1 abrogates Wnt signalling by modulating β-catenin target gene expression and inducing the GSK3-independent degradation of β-catenin. **a**, In Saos2-*TR-E2F1* cells, E2F1 represses c-myc levels without affecting the levels of TCF1 or TCF4. **b**–**d**, E2F1 modulates the expression of endogenous Wnt target genes (qPCR analysis after Tet-induced E2F1 expression at 24 h). **e**, The kinetics of E2F1-induced *AXIN2* and *SIAH1* expression mirrors the E2F1 activation of *CCNE1/Cyclin E* and *CCNA1/Cyclin A*, and precedes E2F1-induced apoptosis. **f**, E2F1 represses the expression of Wnt targets in DLD1 colorectal cancer cells. Levels of mRNA were normalized to GAPDH and the effect of E2F1 is depicted as the ratio between samples after *pCMV-empty* or *pCMV-E2F1* (1 μg each) expression. **g**, **h**, E2F1-induced β-catenin degradation (control expression from the same lysates shown in Supplementary Fig. 4f) is both GSK3- and caspase-independent. Saos2 cells were treated with control (DMSO), GSK3 inhibitors (20 μM SB216763 and 5 mM LiCl), or the caspase inhibitor peptide BOC-aspartyl-FMK (BAF; 100 μM) with or without Tet-induction of E2F1 (western blot). **i**, Co-expression of Bcl-2 (25 ng), pRB (10–25 ng), stabilized tumour-derived β-catenin mutants (10–25 ng), or the GSK3-β inhibitor SB216763 (15 μM), partially rescues E2F1-induced apoptosis. Cell death was depicted as the percentage inhibition of enhanced green fluorescent protein (EGFP) loss at 48 h after transfection with *pCMV-E2F1* or *pCMV-empty* (100 ng each) along with *EGFP* expression construct. All data are expressed as mean ± s.d. ($n = 3$).

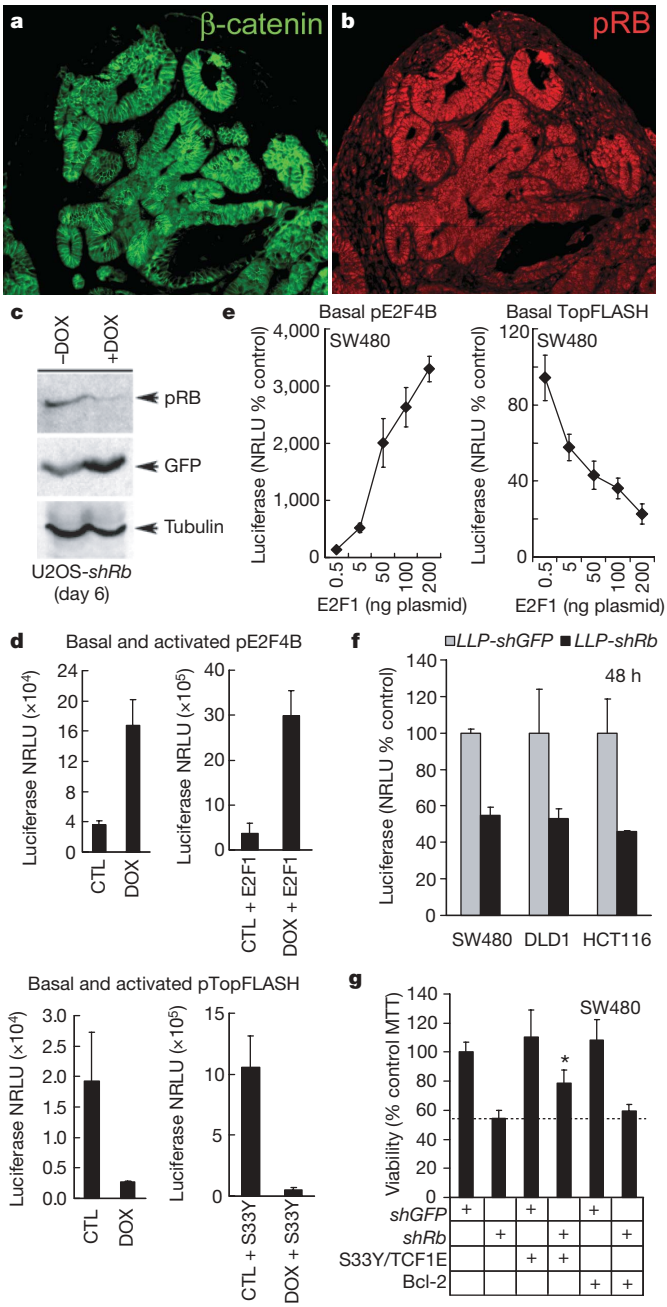


Figure 3 | pRB inactivation abrogates β-catenin/TCF-dependent transcription. **a**, **b**, High levels of pRB and β-catenin co-localize within the tumour epithelium of an *Apc*^{Min} colonic tumour. **c**, Endogenous pRB was depleted for 6 days in U2OS-*shRb* cells (containing DOX-inducible short-hairpin-*Rb* and *GFP* transgenes). **d**, pRB depletion activates E2F and represses TCF activity. Basal and activated E2F and TCF activity was determined by transfection of their respective luciferase reporter plasmids (for 24 h) in the absence (basal) or presence (activated) of E2F1 or S33Y-β-catenin expression constructs. **e**, In SW480 colorectal cancer cells, E2F1 expression is sufficient to activate E2F activity and repress endogenous β-catenin activity in a dose-dependent manner. **f**, Expression of *shRb* (400 ng) is sufficient to repress pTopFLASH activity in three different colorectal cancer cell lines (SW480, DLD1 and HCT116). **g**, pRB inactivation reduces clonogenic survival of SW480 cells and is partly rescued by *S33Y-β-catenin/TCF1E*, but not by *Bcl-2* expression. SW480 cells were Amaxa nucleoporated with either control *LLP-GFP* or *-Rb* silencing constructs (1 μg each) with or without S33Y-β-catenin (250 ng), TCF1E (250 ng) or Bcl-2 (500 ng) expression constructs, and survival was determined at day 5 by MTT assay. All data are expressed as mean ± s.d. ($n = 3$; $^*P < 0.015$ by *t*-test).

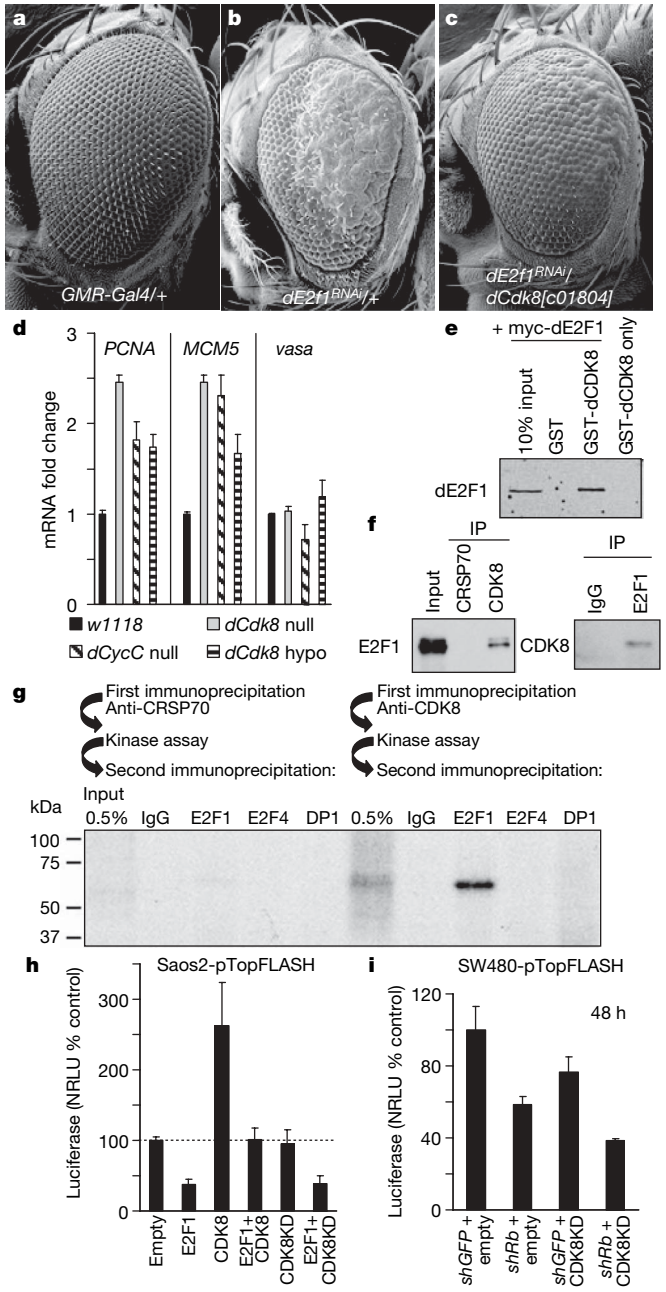


Figure 4 | CDK8 antagonizes E2F1 activity. **a**–**c**, The *dCdk8*^{c01804} mutant dominantly suppresses a rough eye phenotype caused by the eye-specific expression of a *dE2F1*^{RNAi} transgene. **d**, The expression of the dE2F1 target genes *PCNA* and *MCM5* is upregulated in *dCdk8* (null or hypomorphic (hypo) alleles) or *dCycC* null mutant *Drosophila* larvae, whereas the expression of the dE2F2 target gene *vasa* is unaffected. **e**, dCDK8 physically interacts with dE2F1 by GST-pulldown assay. **f**, Co-immunoprecipitation (IP) of human E2F1 and CDK8 from Saos2 whole-cell extracts. As control, E2F1 does not associate with the small-Mediator subunit CRSP70. **g**, CDK8 binds to and specifically phosphorylates E2F1. Kinase-assay was performed after CDK8 or CRSP70 (as control) immunoprecipitation from Saos2 cells. E2F1, E2F4 or DP1 were re-immunoprecipitated and resolved in 12% SDS–polyacrylamide gel electrophoresis (SDS–PAGE). **h**, The expression of wild-type CDK8, but not kinase-dead, abrogates the inhibitory effects of E2F1 on β-catenin/TCF-dependent transcription. pTopFLASH assays were determined at 48 h in co-transfection experiments with 200 ng CDK8 or CDK8KD in Saos2 cells. **i**, Expression of *shRb* (400 ng) and CDK8KD (300 ng) cooperatively repress pTopFLASH activity in SW480 colorectal cancer cells. All data are expressed as mean ± s.d. ($n = 3$).

changes (Supplementary Fig. 6a). In addition, RNAi-mediated depletion of dCDK8, or other components of the CDK8 sub-module, partly suppressed the cell proliferation defects caused by depletion of dE2F1 (Supplementary Fig. 6b, c).

Glutathione *S*-transferase (GST)-pulldown assays demonstrated a strong physical interaction between dCDK8 and dE2F1 (Fig. 4e) that mapped to the dE2F1 transactivation domain (Supplementary Fig. 7). The physical interaction between E2F1 and CDK8 is conserved between species: human E2F1 co-immunoprecipitated with CDK8 from Saos2 cell extracts (Fig. 4f). Moreover, E2F1 was specifically phosphorylated by CDK8 when complexes were incubated in kinase buffer (Fig. 4g). Thus, CDK8 physically interacts with E2F1 and is a conserved negative regulator of E2F1-dependent transcription.

The interaction between CDK8 and E2F1 is particularly notable, as a concurrent study has found significant *CDK8* copy number gains in colorectal cancers: approximately 40% of tumours have copy number gains in both *CDK8* and *RB1* (ref. 1). Chromatin immunoprecipitation (ChIP) experiments on SW480 colorectal cancer cells confirmed that CDK8 and E2F1 are both present at E2F-regulated promoters as well as the *c-MYC* promoter (Supplementary Fig. 8), suggesting an interplay between E2F1, CDK8 and β-catenin/TCF. In Saos2 cells, which contain low basal Wnt activity, expression of CDK8 enhanced β-catenin activation from the pTopFLASH reporter, whereas expression of kinase-dead CDK8 (CDK8KD) had no effect (Fig. 4h). Moreover, expression of CDK8, but not CDK8KD, suppressed the inhibitory effect of E2F1 on the pTopFLASH reporter (Fig. 4h). Similar results were observed for CDK8 expression in HCT116 colorectal cancer cells (Supplementary Fig. 8). Hence, CDK8 and E2F1 have antagonistic effects on β-catenin-mediated transcription, and increasing the levels of CDK8 protects β-catenin from inhibition by E2F1. In agreement with this, the expression of CDK8KD reduced pTopFLASH activity in *APC*-deficient SW480 colorectal cancer cells, and enhanced the inhibition caused by short-hairpin targeting of pRB (Fig. 4i).

Cancer cells acquire multiple mutations during tumorigenesis, and it is a major challenge to explain how each change contributes to malignancy. However, the absence of mutations can also give new insights. It has long been known that colorectal tumour cells fail to mutate *RB1* and typically express elevated levels of this tumour suppressor. The discovery that E2F1 is a potent inhibitor of β-catenin-dependent transcription provides an unexpected and simple explanation to this conundrum. This interaction may also explain why colorectal tumours frequently overexpress c-myc-induced microRNAs that target E2F1 (refs 25 and 26). The discovery by Firestein and colleagues[1] showing significant *RB1* and *CDK8* copy number gains in colorectal cancers is especially intriguing given the evidence that CDK8 is an important modulator of both β-catenin and E2F1. Whereas CDK8 enhances the activity of β-catenin, it represses the activity of E2F1. Consequently, the amplification of *CDK8* may act as a switch, allowing increased β-catenin-dependent transcription that is also resistant to E2F1 inhibition (see model in Supplementary Fig. 1). Reversing this process, such as the inhibition of CDK8 combined with the activation of E2F1, may be useful as a two-pronged strategy to target cancer cells that are driven by deregulated β-catenin activity.

METHODS SUMMARY

Unless otherwise noted, all fly crosses were conducted at 25 °C and phenotypes are depicted in female progeny. Transgenic *dE2f1-dsRNA* (*dE2f1*^{RNAi}) flies were created using a system developed previously[27]. SW480 (A. Burgess), DLD1 (W. Hahn), U2OS-*shRb* (S. Lowe), and Saos2-*TR-E2F1* cells[2] were cultured in DMEM supplemented with 10% fetal bovine serum. *Drosophila* Schneider line 2 (SL2) cells were maintained as previously described[2]. Transient transfections were performed using Fugene-6 reagent (Roche), and in some cases with CellFectin (Invitrogen), according to the manufacturer's instructions. For SW480 survival and DLD1 quantitative PCR (qPCR) experiments, high-efficiency (approximately 70%) gene transfer was accomplished by using Amaxa nucleofection according to the manufacturer's protocol (Program T-020; Kit-T and Kit-L, respectively). All *Drosophila* RNAi in SL2 cells was performed as

described[2], using 50-μg double-stranded RNA (dsRNA) synthesized with T7 RiboMax (Promega) with all conditions normalized using *luciferase*-dsRNA. Luciferase reporter assays were performed as previously described for SL2 cells[2] and mammalian cells[28]. MTT viability assay was performed as previously described[2]. Western blot and immunohistochemical analysis was performed using standard techniques. Dissected third-instar larval disc immunohistochemistry was performed using anti-dE2F1 antibody (T. Orr-Weaver). Immunohistochemical staining of Apc^{Min} mouse tumours was performed as described[19] using anti-Rb (Santa Cruz, sc-50) and β-catenin antibodies (BD Biosciences, 610054). ChIP and data analysis were performed as previously described[29]. Gel shift assays were performed as described[30]. Detailed information on antibodies, fly stocks and plasmids, as well as qPCR, GST-pulldown assay, co-immunoprecipitation and IP-kinase assay conditions, is given in Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 3 March; accepted 24 July 2008.
Published online 14 September 2008.

1. Firestein, R. *et al.* CDK8 is a colorectal cancer oncogene that regulates β-catenin activity. *Nature* doi:10.1038/nature07179 (this issue).
2. Morris, E. J. *et al.* Functional identification of Api5 as a suppressor of E2F-dependent apoptosis *in vivo*. *PLoS Genet* **2,** e196 (2006).
3. Freeman, M. & Bienz, M. EGF receptor/Rolled MAP kinase signalling protects cells against activated Armadillo in the *Drosophila* eye. *EMBO Rep.* **2,** 157–162 (2001).
4. Clevers, H. Wnt/β-catenin signaling in development and disease. *Cell* **127,** 469–480 (2006).
5. Kinzler, K. W. & Vogelstein, B. Lessons from hereditary colorectal cancer. *Cell* **87,** 159–170 (1996).
6. Morin, P. J. *et al.* Activation of β-catenin-Tcf signaling in colon cancer by mutations in β-catenin or APC. *Science* **275,** 1787–1790 (1997).
7. Hallstrom, T. C. & Nevins, J. R. Specificity in the activation and control of transcription factor E2F-dependent apoptosis. *Proc. Natl Acad. Sci. USA* **100,** 10848–10853 (2003).
8. Black, E. P., Hallstrom, T., Dressman, H. K., West, M. & Nevins, J. R. Distinctions in the specificity of E2F function revealed by gene expression signatures. *Proc. Natl Acad. Sci. USA* **102,** 15948–15953 (2005).
9. Sadot, E., Geiger, B., Oren, M. & Ben-Ze'ev, A. Down-regulation of β-catenin by activated p53. *Mol. Cell. Biol.* **21,** 6768–6781 (2001).
10. Rother, K. *et al.* Identification of Tcf-4 as a transcriptional target of p53 signalling. *Oncogene* **23,** 3376–3384 (2004).
11. Ueda, Y. *et al.* p73β, a variant of p73, enhances Wnt/β-catenin signaling in Saos-2 cells. *Biochem. Biophys. Res. Commun.* **283,** 327–333 (2001).
12. Sansom, O. J. *et al. Myc* deletion rescues Apc deficiency in the small intestine. *Nature* **446,** 676–679 (2007).
13. Liu, J. *et al.* Siah-1 mediates a novel β-catenin degradation pathway linking p53 to the adenomatous polyposis coli protein. *Mol. Cell* **7,** 927–936 (2001).
14. Matsuzawa, S. I. & Reed, J. C. Siah-1, SIP, and Ebi collaborate in a novel pathway for β-catenin degradation linked to p53 responses. *Mol. Cell* **7,** 915–926 (2001).
15. Hughes, T. A. & Brady, H. J. E2F1 up-regulates the expression of the tumour suppressor axin2 both by activation of transcription and by mRNA stabilisation. *Biochem. Biophys. Res. Commun.* **329,** 1267–1274 (2005).
16. Hallstrom, T. C., Mori, S. & Nevins, J. R. An E2F1-dependent gene expression program that determines the balance between proliferation and cell death. *Cancer Cell* **13,** 11–22 (2008).
17. Gope, R. *et al.* Increased expression of the retinoblastoma gene in human colorectal carcinomas relative to normal colonic mucosa. *J. Natl Cancer Inst.* **82,** 310–314 (1990).
18. Kucherlapati, M. H., Nguyen, A. A., Bronson, R. T. & Kucherlapati, R. S. Inactivation of conditional Rb by Villin-Cre leads to aggressive tumors outside the gastrointestinal tract. *Cancer Res.* **66,** 3576–3583 (2006).
19. Haigis, K., Sage, J., Glickman, J., Shafer, S. & Jacks, T. The related retinoblastoma (pRb) and p130 proteins cooperate to regulate homeostasis in the intestinal epithelium. *J. Biol. Chem.* **281,** 638–647 (2006).
20. Williams, J. P. *et al.* The retinoblastoma protein is required for Ras-induced oncogenic transformation. *Mol. Cell. Biol.* **26,** 1170–1182 (2006).
21. Dickins, R. A. *et al.* Probing tumor phenotypes using stable and regulated synthetic microRNA precursors. *Nature Genet.* **37,** 1289–1295 (2005).
22. Faux, M. C. *et al.* Restoration of full-length adenomatous polyposis coli (APC) protein in a colon cancer cell line enhances cell adhesion. *J. Cell Sci.* **117,** 427–439 (2004).
23. Kim, S., Xu, X., Hecht, A. & Boyer, T. G. Mediator is a transducer of Wnt/β-catenin signaling. *J. Biol. Chem.* **281,** 14066–14075 (2006).
24. Malik, S. & Roeder, R. G. Dynamic regulation of pol II transcription by the mammalian Mediator complex. *Trends Biochem. Sci.* **30,** 256–263 (2005).
25. He, L. *et al.* A microRNA polycistron as a potential human oncogene. *Nature* **435,** 828–833 (2005).
26. O'Donnell, K. A., Wentzel, E. A., Zeller, K. I., Dang, C. V. & Mendell, J. T. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* **435,** 839–843 (2005).
27. Lee, Y. S. & Carthew, R. W. Making a better RNAi vector for *Drosophila*: use of intron spacers. *Methods* **30,** 322–329 (2003).
28. Dick, F. A., Sailhamer, E. & Dyson, N. J. Mutagenesis of the pRB pocket reveals that cell cycle arrest functions are separable from binding to viral oncoproteins. *Mol. Cell. Biol.* **20,** 3715–3727 (2000).
29. Di Stefano, L., Jensen, M. R. & Helin, K. E2F7, a novel E2F featuring DP-independent repression of a subset of E2F-regulated genes. *EMBO J.* **22,** 6289–6298 (2003).
30. Hurford, R. K. Jr, Cobrinik, D., Lee, M. H. & Dyson, N. pRB and p107/p130 are required for the regulated expression of different sets of E2F responsive genes. *Genes Dev.* **11,** 1447–1463 (1997).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank many investigators for their gifts of cell lines, plasmids and fly stocks, especially S. Artavanis-Tsakonas. We thank D. Rennie and the Massachusetts General Hospital Cutaneous Biology Research Center Transgenic Fly Core for embryo injections, and B. Fowle for his help with SEM imaging. We thank A. McClatchey, J. Settleman, C. Seum and T. Orr-Weaver for their gifts of antibodies. We thank our colleagues at the Massachusetts General Hospital (MGH) Cancer Center for discussions. E.J.M. and J.-Y.J. are supported in part by a Ruth L. Kirschstein Award and a Tosteson Postdoctoral Fellowship, respectively. L.D.S. is supported by the MGH ECOR Fund for Medical Discovery. N.-S.M. is a Leukemia and Lymphoma Society Special Fellow. K.M.H. was supported by a Career Development award from the Harvard Gastrointestinal Specialized Program of Research Excellence (GI-SPORE) (P50-CA127003). N.J.D. was supported by a scholarship from the Saltonstall Foundation. This study was supported by grants from the National Institutes of Health to N.J.D. (GM81607, GM053203) and A.M.N. (GM071449).

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to N.J.D. (dyson@helix.mgh.harvard.edu).

METHODS

Fly stocks, transgenes and genetic crosses. The following stocks were used for these studies: *GMR-Gal4*, *ptc-Gal4*, *UAS-arm* (*S2*; *4783*); *UAS-dTCF/pan* (*4837*); *UAS-dTCFΔN/pan* (*4784*); *UAS-GSK3/sgg* (*5361*) (Bloomington Stock Center); *GMR-arm** (*Y55*; M. Bienz); $dCdk8^{K185}$ (null) and $dCycC^{Y5}$ (null) (H.-M. Bourbon); *GMR-wIR* (R. Carthew); and *PCNA-GFP* (B. Duronio). The $dE2f1^{RNAi}$ primer sequence (sub-cloned inverted into the pWIZ vector) was 5′-TTATTTCAAACGCCCTACCG-3′ and 5′-GAATTGCATCTGCAGTGAGC-3′ (verified by sequencing). All transgenic fly embryo injections were performed by the MGH Cutaneous Biology Research Center Transgenic Fly Core. Approximately 30 different transgenic lines carrying one or multiple transgenes were balanced and recombined with different *Gal4* lines using standard *Drosophila* genetic methods. Because the $dE2f1^{RNAi}$ phenotypes in both the *Drosophila* eye (*GMR-Gal4,UAS-*$dE2f1^{RNAi}$ *#10*) and wing (*ptc-Gal4,UAS-*$dE2f1^{RNAi}$ *#3*) could be modified by known factors that interact with the RB-E2F pathway (J.-Y.J. and N.J.D., unpublished observations), we designed and performed a dominant modifier genetic screen. We screened approximately 6,500 *piggyBac* transposon insertion lines of the Exelixis mutant collection (J.-Y.J., A.H. and N.J.D., unpublished results) that was maintained at the Harvard *Drosophila* Collection (S. Artavanis-Tsakonas). A detailed description of the genetic screen will be presented elsewhere.

Cell culture and gene transfer. Plasmids used included *pCMV-E2F1*, *pCMV-DP1*, *pCMV-E2F2*, *pCMV-E2F3*, *pCMV-E2F4*, *pCMV-E2F1ΔX*, *pCMV-E2F1-121*, *pCMV-E2F1-143*, *pCMV-E2F1-144*, *pGST–E2F1*, *pGST–E2F4*, *pGST–DP1* (K. Helin); *pSFFV-p53* (S. Korsmeyer); *pCMV-neoBam*, *pCMV-β-gal4*, *pCMV-pRB*, *pE2F4B* (F. Dick); *pCMV-DNDP1* (*Δ103–126*) (M. Classon); *8×SuperTop*, *8×SuperFop*, *pCMV-CDK8*, *pCMV-CDK8KD*, *pLKO-shCDK8-1*, *pLKO-shGFP* (W. Hahn); *pCMV-E2F1/3 chimaeras* (J. Nevins); *pcDNA3-DNTCF1*, *pcDNA3-DNTCF4* (H. Clevers); *pEVR-TCF1E* (M. Waterman); *pGL3OT*, *pGL3OF* (E. Hay); *pEGFP-N1* (BD Biosciences); *pLPC-empty*, *pLPC-TAp73*, *pClneo-Δ45-β-catenin*, *pCIneo-S33Y-β-catenin*, *pCIneo-β-catenin*, *LLP-shGFP* and $LLP\text{-}shRb^{CD}$ (targeting sequence GGTTGTGTCGAAATTGGATCA) (J. Rocco). Inhibitors and chemical reagents include LiCl GSK3 inhibitor (Fisher, 121-500), SB216763 GSK3-β inhibitor (Sigma, S3442), BOC-Aspartyl-FMK caspase inhibitor (Enzyme Systems Products, FK-011), MG101 (Sigma, A6185) and MG132 (Sigma, C2211).

Luciferase reporter and EGFP viability assays. For pTopFLASH assays, Saos2 cells were transfected in six-well plates with 100–200 ng *pTopFLASH* reporter plus 100 ng S33Y-β-catenin expression construct, along with 100 ng *pCMV-β-gal* as normalization control after β-gal assay. Unless otherwise specified, luciferase assays were performed 48 h after transfection (data are expressed as mean ± s.d., $n = 3$). The S33Y-β-catenin construct was omitted for *pTopFLASH* assays in colorectal cancer cells. For EGFP co-transfection viability assay, 100 ng of *pEGFP-N1* was co-transfected with the indicated plasmids. Whole-cell lysates were prepared with GFP homogenizing buffer and assayed fluorometrically 48 h after transfection (excitation/emission $\lambda = 488/511$ nm). Tet-induced survival experiments were done under low serum (0.5%) conditions.

Antibodies for western and immunohistochemical analyses. Other antibodies used included dE2F1 (polyclonal anti-rabbit, C. Seum), E2F1 (Santa Cruz, sc-193), E2F4 (Santa Cruz, sc-1082), DP1 (Santa Cruz, sc-610), CRSP70 (Santa Cruz, sc-9426), c-myc (Santa Cruz, sc-40), TCF1 (Santa Cruz, sc-8589), TCF4 (Santa Cruz, sc-8632; Upstate, 05-511 for ChIP), Cyclin E (Santa Cruz, sc-247), Cyclin A (Santa Cruz, sc-596), pan-MAPK (BD Biosciences, 612641), β-catenin (Cell Signaling, 9562), CDK8 (Santa Cruz, sc-1521), pRB (Santa Cruz, sc-50), Arm (N27A1; Developmental Studies Hybridoma Bank), β1-integrin (BD Biosciences, 610467), E-cadherin (BD Biosciences, 610181), N-cadherin (BD Biosciences, 610920), α-catenin (BD Biosciences 610193), FAK (BD Biosciences 610087), phospho-FAK Tyr397 (BD Biosciences, 611806), GSK3 (Cell Signaling, 9315), phospho-GSK3 (Upstate, 05-413), Src (Cell Signaling, 2108), phospho-Src Tyr416 (Cell Signaling, 2101), STAT3 (Cell Signaling, 9132), phospho-STAT3 Tyr705 (Cell Signaling, 9131), phospho-STAT3 Ser727 (Cell Signaling, 9134), STAT5 (Cell Signaling, 9310), phospho-STAT5 Tyr694 (Cell Signaling, 9356), SHC (BD Biosciences, 610081), anti-HA-epitope (Sigma, H6908), and GFP (Sigma, G1544).

Real-time quantitative PCR. Total RNA was prepared using RNeasy Extraction Kit (Qiagen). Reverse transcription PCR (RT–PCR) was performed using Taq Man Reverse Transcription (PE Applied Biosystems) according to the manufacturer's specification. RT–PCR was performed using an ABI prism 7900 HD Sequence Detection system. Relative mRNA levels were determined using the SYBR Green I detection chemistry system (Applied Biosystems). Quantification was performed using the comparative C_T method as described in the manufacturer's manual, and *GAPDH* or *18S rRNA* was used as normalization control. All primers were designed with Primer Express 1.0 software (Applied Biosystems) following the manufacturer's suggested conditions.

Forward and reverse primer sequences. These included: E2F1-f (CATCCCTCACCACAGATCCC), E2F1-r (AACAGCGGTTCTTGCTCCAG); c-MYC-f (CGGATTCTCTGCTCTCCTCG), c-MYC-r (CCACAGAAACAACA TCGATTTCTT); AXIN1-f (ACAGCATCGTTGTGGCGTACT), AXIN1-r (CACAGTCAAACTCGTCGCTCAC); AXIN2-f (ATTCGGCCACTGTTCAGA CG), AXIN2-r (GACAACCAACTCACTGGCCTG); SIAH1-f (ATGGTCATA GGCGACGATTGAC), SIAH1-r (AAGCTGTGCAATGCTGGTGTC); CD44-f (CTGCAAGGCTTTCAATAGCACC), CD44-r (CGTGCCCTTCTATGAACC CAT); PPARδ-f (CTTCCACTACGGTGTTCATGCA), PPARδ-r (GCACTTG TTGCGGTTCTTCTTC); CCNE1-f (TGCAGAGCTGTTGGATCTCTGTG), CCNE1-r (GGCCGAAGCAGCAAGTATACC); CDC25-f (TGAATACGAGGG AGGCCACATCAA), CDC25-r (ACACGCTTGCCATCAGTAGGTACA); TP73-f (CGTACTCCCCGCTCTTGAAG), TP73-r (TCCGCTTTCTTGTAAACAGGC); GAPDH-f (CATGTTCGTCATGGGTGTGAACCA), GAPDH-r (GTGATGGCAT GGACTGTGGTCAT).

Chromatin immunoprecipitation. Chromatin extracts were incubated at 4 °C overnight with antibodies specific for E2F1, CDK8 and anti-HA antibody as control. Immunocomplexes were recovered with protein A and G Sepharose beads. DNA was recovered and dissolved in 150 μl of water. RT–qPCR was performed as described above.

Forward and reverse primer sequences. These included: P107-f (GGCCAAGGACAGGTCTTTCAG), P107-r (AAAGACGCCCAGAGATGCA GC); CCNA1-f (CTGCTCAGTTTCCTTTGGTTTAC), CCNA1-r (AAAGACG CCCAGAGATGCAGC); ACYLCOA-f (CCTTCATTGGGATCACCACG), ACYLCOA-r (GGAGATGAGTACCAGCAGGTTG); MYC-EBE-f (TCCGCC TGCGATGATTTATAC), MYC-EBE-r (CAGAGTAAGAGAGCCGCATGAA); MYC-TBE1-f (TCTCCCGTCTAGCACCTTTGA), MYC-TBE1-r (CACGGAG TTCCCAATTTCTCAG); MYC-TBE2-f (GCTCTCCAAGTATACGTGGCAA), MYC-TBE2-r (TCAGAGCGTGGGATGTTAGTGTA).

GST-pulldown assay. GST–fusion proteins were expressed in *Escherichia coli* BL21 cells and purified from the lysates by glutathione sepharose beads (Pharmacia) with extensive washing. The amount of bead-bound GST protein was determined by SDS–PAGE and Coomassie blue staining and normalized. The dE2F1 fragments, shown in Supplementary Fig. 7, were subcloned into *pGEX-2TKN* vector. To make Myc-tagged dE2F1 and HA-tagged dCDK8, full-length cDNAs were subcloned into *pcDNA4TO* or *pcDNA3HA*, respectively. Each construct was expressed in 293T cells and extracted in binding buffer (20 mM HEPES pH 7.6, 140 mM NaCl, 0.1 mM EDTA, 10% glycerol, 1 mM dithiothreitol (DTT), 1 mM benzamidine, 0.2 mM phenylmethylsulphonyl fluoride (PMSF), and 1 μg ml^{-1} aprotinin) with 0.5% NP-40. The whole-cell extracts were diluted fourfold with binding buffer without detergent. S2 cell extracts were prepared similarly. Whole-cell extract was applied to 50 μl of GST–fusion protein beads and incubated at 4 °C for 3 h. Beads were washed seven times with 1 ml wash buffer containing 50 mM Tris-HCl at pH 8.0, 250 mM KCl, 0.1 mM EDTA, 10% glycerol, 0.1% NP-40, 1 mM DTT, 1 mM benzamidine, 0.2 mM PMSF and 1 μg ml^{-1} aprotinin. The interacting proteins were then eluted with two bead volumes of binding buffer plus 0.3% sarkosyl at 4 °C for 1 h.

Co-immunoprecipitation. Saos2-*TR-E2F1* cells were collected and rinsed with PBS after culture in the presence of 0.2 μg ml^{-1} of tetracycline for 12 h. Cell pellets were resuspended in immunoprecipitation buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.1 mM EDTA, 10% glycerol, 0.1% NP-40, 1 mM DTT, 0.2 mM PMSF, 1 mM benzamidine and 1 μg ml^{-1} aprotinin). Whole-cell extract was prepared by passing through 20(1/2)G needle for five times followed by incubation on ice for 15 min and centrifugation at 12,000*g* for 20 min. For immunoprecipitation, whole-cell extracts were incubated with 30 μl of protein-G beads pre-coated with antibodies against CDK8 (sc-1521), E2F1 (sc-193), control normal rabbit serum (IgG) or CRSP70 (goat anti-human CRSP70, Santa Cruz sc-9426). After rotating at 4 °C for 2 h, the beads were washed five times with 1 ml immunoprecipitation buffer. The antigen-associated proteins were eluted with 0.3% sarkosyl in immunoprecipitation buffer, resolved on 10% SDS–PAGE, and visualized using standard western analysis.

Immunoprecipitation-kinase assay. CDK8 or control (CRSP70) proteins were immunoprecipitated from whole-cell extracts of Saos2-*TR-E2F1* cells as described in the co-immunoprecipitation section. The beads were washed twice with 1 ml immunoprecipitation buffer followed by two washes with 1 ml kinase assay buffer (50 mM Tris-HCl pH 7.4, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM $MnCl_2$, 10% glycerol, 0.2 mM PMSF, 5 mM DTT and 1 μg ml^{-1} aprotinin). For kinase assays, the resulting beads were incubated in 50 μl of kinase assay buffer containing 20 μCi of [γ-^{32}P]ATP at room temperature for 30 min. The reaction was terminated and the kinase-associated proteins were eluted with 0.3% sarkosyl in immunoprecipitation buffer. After dilution ten times with

immunoprecipitation buffer, an equal amount of the supernatant was applied to protein-A beads pre-coated with antibodies against IgG (normal rabbit serum), E2F1 (sc-193), E2F4 (sc-1082) or DP1 (sc-610). The mixture was rocked at 4 °C for 2 h and the beads were washed four times with 1 ml immunoprecipitation buffer. The samples were boiled, resolved in 12% SDS–PAGE and then transferred to a PVDF membrane. The *in vitro* phosphorylated proteins were visualized by autoradiography.

Electrophoretic mobility gel shift assay. Briefly, 2 μg of nuclear extracts were incubated with ^{32}P-labelled double-stranded oligonucleotides containing E2F or TCF sites site (less than 10 pg) for 30 min at 4 °C in binding buffer. For supershift experiments, antibodies were pre-incubated for 30 min. Samples were then loaded on a 4% polyacrylamide gel and visualized by autoradiography.

Double-strand oligonucleotide probe sequences. These included: MYC-wtE2F (5′-GCTTCTCAGAGGCTTGGCGGGAAAAAGAA); MYC-mtE2F (5′-GCTTCTCAGAGGCTTGTAGGGAAAAAGAA); MYC-TBE1 (5′-GTCTAGCACCTTTGATTTCTCCCAAACCC); MYC-TBE2 (5′-CTGGGACTCTTGATCAAAGCGCGGCCCTT); TOP (5′-GCCCCCTTTGATCTTACCCCCTTTGATCT); Canonical wtE2F (5′-ATTTAAGTTTCGCGCCCTTTCTCAAATTT); Canonical mtE2F (5′-ATTTAAGTTTCGATCCCTTTCTCAAATTT).

LETTERS

Replication fork movement sets chromatin loop size and origin choice in mammalian cells

Sylvain Courbet[1], Sophie Gay[1]*, Nausica Arnoult[1]*, Gerd Wronka[1], Mauro Anglana[1], Olivier Brison[1] & Michelle Debatisse[1]

Genome stability requires one, and only one, DNA duplication at each S phase. The mechanisms preventing origin firing on newly replicated DNA are well documented[1], but much less is known about the mechanisms controlling the spacing of initiation events[2,3], namely the completion of DNA replication. Here we show that origin use in Chinese hamster cells depends on both the movement of the replication forks and the organization of chromatin loops. We found that slowing the replication speed triggers the recruitment of latent origins within minutes, allowing the completion of S phase in a timely fashion. When slowly replicating cells are shifted to conditions of fast fork progression, although the decrease in the overall number of active origins occurs within 2 h, the cells still have to go through a complete cell cycle before the efficiency specific to each origin is restored. We observed a strict correlation between replication speed during a given S phase and the size of chromatin loops in the next G1 phase. Furthermore, we found that origins located at or near sites of anchorage of chromatin loops in G1 are activated preferentially in the following S phase. These data suggest a mechanism of origin programming in which replication speed determines the spacing of anchorage regions of chromatin loops, that, in turn, controls the choice of initiation sites.

We previously used DNA combing to analyse the replication dynamics along the *AMPD2* (adenosine monophosphate deaminase 2) region[3] in Chinese hamster fibroblasts highly amplified for this locus (Supplementary Fig. 2). Mutant lines were selected for their resistance to coformycin, an inhibitor of AMPD2, which depletes cells of guanine derivatives[4]. In mutant line 471, replication forks proceed at about 1.3 kilobases (kb) min^{-1} (referred to below as 'fast'). The density of initiation events is low, and about 80% of the events observed within the locus take place at *oriGNAI3*, a previously characterized replication origin[5–7], the remaining events being scattered between five other discrete origins. In contrast, forks travel at about 0.6 kb min^{-1} in line 474 (referred to below as 'slow'). Initiation events occur at high density and seem evenly distributed between the five origins and *oriGNAI3* (ref. 3). Because the selection of mutant lines relied on unbalanced nucleotide pools, we reasoned that abnormal pools of DNA precursors might persist in cells of line 474 ('474 cells'). Indeed, the simple addition of adenine and uridine to the culture medium (referred to below as A + U) was sufficient to convert the replication pattern of those slow cells to that observed in fast cells[3] (Fig. 1a). Conversely, hydroxyurea-induced nucleotide starvation slows fork speed in cells of line 471 ('471 cells') and results in the redistribution of initiation between all six origins of the locus. Altogether, these results reveal a highly controlled switch in origin usage when cells face variations in growth conditions[2,3], offering a model with which to address the question of how and when origin choice is established.

We first determined whether the initiation pattern is controlled by nucleotide pools or by fork movement. Fast 471 cells were treated with aphidicolin (75 ng ml^{-1}), a DNA polymerase inhibitor that does not interfere with the pools. Analysis of combed DNA molecules

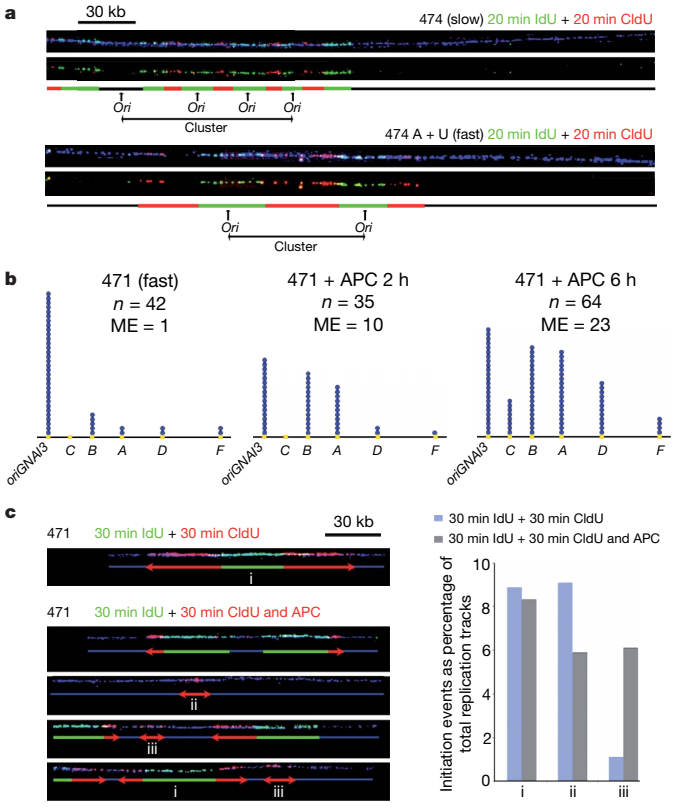


Figure 1 | The density of active origins depends on replication speed. **a**, Combed DNA molecules (DNA stained in blue) from 474 cells grown in normal or in A + U medium. A schematic representation of the replication patterns is shown below each molecule. **b**, Localization of initiation events (one dot per event) along the *AMPD2* region in 471 cells grown in normal medium or challenged for 2 h or 6 h with 75 ng ml^{-1} aphidicolin (APC). ME, number of molecules with multiple initiation events; *n*, number of molecules analysed. **c**, DNA molecules (as in **a**) from 471 cells labelled with IdU, then with CldU with or without APC (75 ng ml^{-1}), as described in ref. 27. Left: combed molecules. Initiation events occurring during: the first pulse (i), the second pulse on DNA regions that were not replicating during the first pulse (ii), and the second pulse on DNA regions that did start to replicate during the first pulse (iii). Right: percentage of events observed in each category; $n = 450$.

[1]Institut Curie, 26 rue d'Ulm, 75248 Paris, France; UPMC Univ. Paris 06, F-75005 Paris, France; CNRS UMR 7147.
*These authors contributed equally to this work.

from cells labelled with iododeoxyuridine (IdU) then with chlorodeoxyuridine (CldU) showed that replication speed decreased to about 0.4 kb min^{-1} but cell cycle progression remained roughly normal under these conditions (Supplementary Fig. 3). The distribution of initiation events along the AMPD2 locus was studied by combining fluorescence *in situ* hybridization (FISH) with detection of IdU and CldU on newly synthesized DNA[3]. We observed a strong decrease in *oriGNAI3* prominence 2 h after the addition of aphidicolin, and initiation events seemed evenly distributed between the different origins after 6 h. Loci showing more than one initiation event increased correspondingly (Fig. 1b). Note that *oriGNAI3* still fired, at least on half the molecules that we observed (Supplementary Fig. 4a); loss of its prominence therefore results partly from an increased efficiency of the other origins.

We also performed a global analysis of the replication dynamics in 471 cells that were labelled with IdU in normal medium (fast) and then with CldU with or without aphidicolin. Comparison of the lengths of the IdU and CldU tracks showed that fork speed decreased almost instantaneously on the addition of aphidicolin (Supplementary Fig. 4b). The distribution of initiation events occurring during each labelling period was studied by focusing on 200-kb DNA regions (Fig. 1c). As expected, the same frequency of initiation was observed during the first pulse whether or not the cells were treated with aphidicolin during the second pulse (Fig. 1c, i). In cells not challenged by aphidicolin, initiation during the CldU-labelling period was observed with one-tenth the frequency on regions already labelled by IdU than on regions free of IdU signal (Fig. 1c, compare ii with iii). This confirms that the density of initiation events is low when replication forks travel fast. By contrast, in cells treated with aphidicolin, the probability of firing during the second pulse was similar on both types of region (Fig. 1c, compare ii with iii), confirming that the density of initiation events is high when forks travel slowly.

Taken together, our results show that fork speed itself, rather than the nucleotide pools, controls the pattern of initiation. They further show that, within 30 min, the cells start to compensate for the decrease in fork speed by mobilizing latent origins, which are thus able to change their fate within S-phase. Because licensing is prevented after the onset of S phase[8], we conclude that the pre-initiation complex is present on these latent origins.

We then studied the reciprocal situation, namely the specific silencing of some origins when 474 cells are shifted from slow to fast conditions. A global analysis of the replication dynamics showed that fork speed reached the maximum value (about 1.8 kb min^{-1}) 15 min after A + U addition (Supplementary Fig. 5a). To assess the distribution of initiation events, we focused our analysis on long DNA molecules (Fig. 1a). After only 2 h in A + U, the density of initiation events decreased by 30% and remained stable at later time points (Supplementary Fig. 5b). At each time point, initiation events seemed confined to a restricted region along any individual molecule that we observed (Fig. 1a). This is consistent with previous work in yeast and mammalian cells showing that origins are functionally organized as clusters[9–11]. Strikingly, the total length of the labelled regions relative to that of the DNA molecules was constant (Supplementary Fig. 5b). Hence, modulation of the density of initiation according to replication speed occurs mostly at the level of individual clusters, a finding consistent with the results obtained in aphidicolin-treated 471 cells (Fig. 1c).

We also determined how the replication pattern evolved along the *AMPD2* locus after shifting 474 cells to A + U medium (Supplementary Fig. 6). We found that the proportion of loci displaying close initiation events (separated by less than 90 kb) decreased from 80% to about 50% within 6 h. However, it took about 16 h before *oriGNAI3* prominence was established, roughly coinciding with the doubling time of these cells. Because we observed cells in only a narrow window of S phase (the locus replication time), this correlation suggested that the cells have to go through a complete cell cycle in fast medium to reprogram the relative efficiency of the origins. We tested this hypothesis by selecting mitotic 474 cells (slow) and then replating them in fast conditions. We specifically analysed cells at the first and second S phases after replating (Fig. 2). In both cases, closely spaced initiation events were observed on about 50% of the molecules displaying several events, as in unsynchronized fast 474 cells. The initiation events remained evenly distributed between all the origins of the locus during the first S phase, and *oriGNAI3* became prominent at the second S phase.

We conclude that remodelling of the initiation pattern by an increased speed occurs in two steps. First, some origins are randomly silenced, probably because they are passively replicated by forks emanating from origins activated earlier. Next, if conditions of high replication speed are maintained, an additional mechanism commits some origins, such as *oriGNAI3*, to fire preferentially during the following S phase. This origin hierarchy remains stable in the absence of further speed variations[3].

Replicon size, which is dictated by the spacing of active origins, has long been correlated with the length of chromatin loops[12]. We therefore

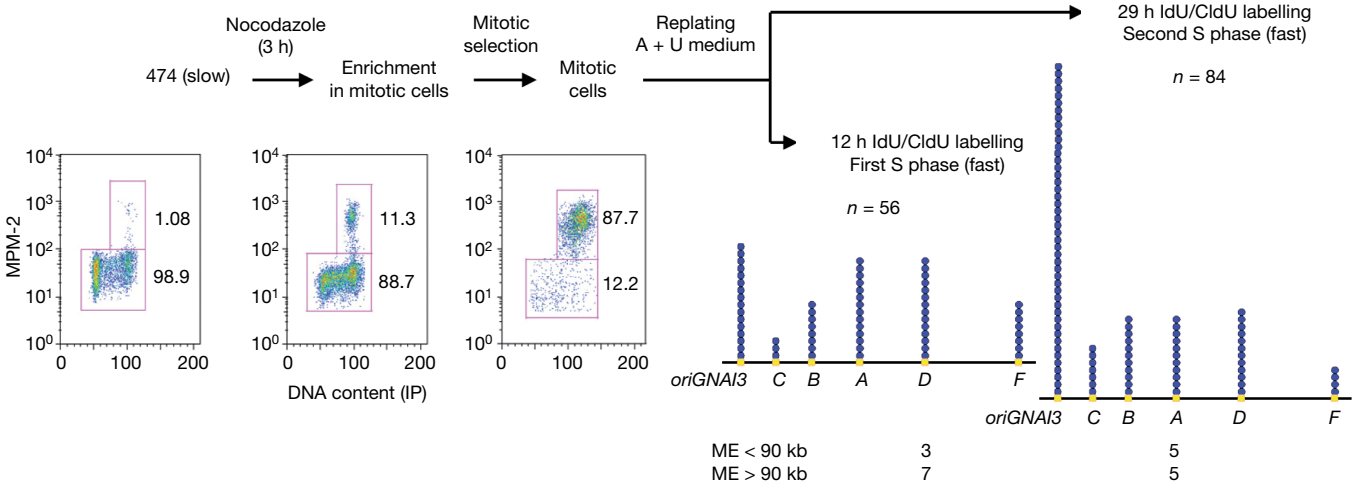


Figure 2 | *oriGNAI3* prominence appears only during the second S phase after shifting 474 cells to fast conditions. Left: FACS analysis of cells growing exponentially, treated with nocodazole, and recovered after mitotic shake-off. The lower box contains cells in G1 + S + G2; the upper box contains mitotic cells. The percentage of cells in each group is indicated. Right: distribution of initiation events along the *AMPD2* locus during the first and the second S phases after replating in A + U medium (symbols are as in Fig. 1b). Below, multiple initiation events (ME) are separated into two categories depending on the inter-origin distance.

studied the size of the loops surrounding the nuclear matrix (Supplementary Fig. 7) and found that they were periodically remodelled during the cell cycle. Furthermore, in G1 nuclei, their mean size increased twofold in fast 474 cells in comparison with slow ones. To some extent, the correlation held in S and G2 phases, even though the difference was less striking. The study of fast and aphidicolin-treated 471 cells confirmed these observations (Supplementary Fig. 8), establishing a correlation between fork speed and loop size. The existence of the nuclear matrix remains controversial[13]. However, a functional relationship between replication and attachment to an operationally defined matrix is now well documented[14]. For example, in Chinese hamster ovary (CHO) cells, a matrix attachment region (MAR) is required for the maintenance of plasmids not integrated in the chromosomes[15]. Moreover, all the origins of the *AMPD2* locus co-localize with MARs[16], similarly to many origins characterized in vertebrates[17]. Here we further show that, depending on the growth conditions, halos with different and specific sizes could be obtained from the same cells, salt-extracted and further treated in parallel.

We then determined whether the spatial distribution in the halos of *oriGNAI3*, *oriA*, *oriB* and a non-origin sequence depended on the replication dynamics (Fig. 3a). In small halos seen with G1 nuclei of slow 474 cells, all four sequences were similarly distributed between the matrix and the loops. In large halos obtained with G1 nuclei of fast 474 cells, whereas *oriB*, *oriA* and the non-origin sequence still seemed distributed, *oriGNAI3* localized preferentially at the matrix. In G2 nuclei of these cells, *oriGNAI3*, like *oriB*, was randomly distributed in the halos (Fig. 3a). The very same distribution of *oriGNAI3* was observed in fast 471 cells (Supplementary Fig. 8a). This origin therefore strikingly relocalizes from the matrix to the loops between G1 and G2 and from the loops to the matrix during mitosis or early G1 in cells grown in fast conditions. By studying cells sorted by fluorescence-activated cell sorting (FACS), we found that *oriGNAI3* relocalized from the matrix to the loops during S phase (Fig. 3a and Supplementary Fig. 7e). This agrees with previous studies showing that newly synthesized DNA moves away from the base of the loops in mammalian cells[18] and that origins are pulled out of the replisome on fork progression in yeast cells[19]. We then determined when *oriGNAI3* relocalized from the loops to the matrix by performing a time-course analysis of *oriGNAI3* and *oriB* in fast 474 cells released from a mitotic block (Fig. 3b). *OriGNAI3* was reattached to the matrix as early as 1 h after release, suggesting that the switch occurs in mitosis. This interpretation is supported by previous data showing that origins can be reset in mitotic *Xenopus* egg extracts in a topo-II-dependent manner[20,21] and that resetting correlates with significant remodelling of chromatin loops[21]. In addition, specific movements of DNA sequences during mitosis were observed *in vivo* in CHO cells[22]. Our data do not account directly for the observation that initiation sites are determined in G1, at a step called the origin decision point[23,24]. However, the attachment of origins to the matrix may be a prerequisite for the further selection of initiation sites.

To determine whether the pattern of initiation during the previous S phase sets loop size and origin localization in G1, we studied the fate of mitotic 474 cells (slow) that had been replated in either slow or fast conditions (Fig. 3c). At the first G1 phase after the shift to fast conditions, loop size and both *oriGNAI3* and *oriB* localization remained the same as in slow cells. In contrast, at the second G1 phase, loop size increased and *oriGNAI3* localized preferentially to the matrix. The study of the reciprocal situation, namely the fate of mitotic 471 cells (fast) replated in the presence of aphidicolin, confirmed that halos seemed remodelled only at the second G1 phase after the shift (Supplementary Fig. 8b). We conclude that the structural organization of the loops in G1 depends on the replication pattern in the previous S phase, and that the association of each origin with the matrix probably determines its probability of firing.

To account for the observation that the size of the loops in G1 depends on the replication dynamics during the previous S phase, we propose that replication marks, whose spacing depends on the density of initiation events, are present on the DNA until mitosis. We infer that such marks could be termination regions. Under fast conditions, large replicons would lead to distant marks that, in turn, organize large loops favouring the anchorage of origins with the highest affinity for the matrix. In slow conditions, the close DNA marks would give rise to small loops that offer all potential origins an opportunity to bind to the matrix. Finally, we postulate specifically that attached origins fire earlier than unattached ones, thereby establishing the flexible origin use observed in our experiments (Supplementary Fig. 1).

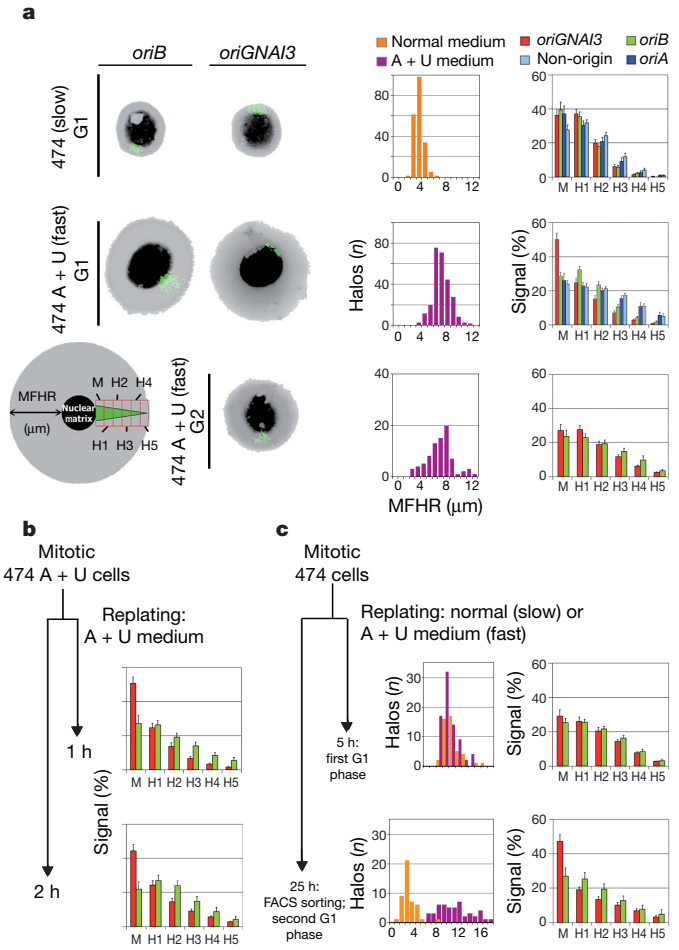


Figure 3 | Large halos and preferential binding of *oriGNAI3* to the nuclear matrix in G1 of 'fast' 474 cells. a, Exponentially growing cells at steady state. Left: halos from cells in G1 and G2. FISH was performed with probes for *oriGNAI3* or *oriB* (green signals). DNA is shown in grey (reverse DAPI). MFHR, maximum fluorescence halo radius. The position of the stacks used for quantification of the FISH signal on all halos is shown: M is positioned at the halo-matrix border (0 to 1.5 μm from the matrix) and H1 to H5 are positioned every 1.5 μm. Right: histograms showing the MFHR and the percentage of FISH signals per stack. The non-origin sequence maps between *oriD* and *oriF*. **b**, Synchronized fast 474 cells in early G1. Histograms as in **a**. **c**, Synchronized 474 cells shifted from slow to fast conditions. Histograms as in **a**. Because the cells were partly desynchronized in the second G1 phase, they were sorted according to their DNA content. All error bars represent the s.e.m. of 35–50 halos.

METHODS SUMMARY

Preparation of halos, and FISH. Halos were prepared as described[12,21], with the following changes: nuclei were isolated by incubating cells in ice-cold NP40 buffer (0.5% Nonidet P40, 10 mM $MgCl_2$, 0.5 mM $CaCl_2$, 25 mM Tris-HCl pH 8.0) for 5 min. Nuclei were washed in ice-cold PBS, stained with 4,6-diamidino-2-phenylindole (DAPI; 2 μg ml^{-1}) and sorted with a FACSVantage (Becton Dickinson) according to their DNA content. Sorted nuclei were spread on SuperFrost slides by Cytospin centrifugation before salt extractions as

described[21]. Halos were then fixed for 10 min with 2% formaldehyde. FISH experiments were performed as described[25]. Maximum fluorescence halo radius was measured with Fluovision software (Imstar). Quantification of FISH signals was performed with ImageQuant v. 5.0 software (FujiFilm).

Cell synchronization procedure. Mitotic selection was performed as described[26], after treatment of the cells for 3 h with nocodazole (200 nM). The quality of the mitotic cell preparations was verified by FACS analysis of MPM2-positive (mitotic protein monoclonal 2) cells. The cells were replated in fresh medium and then recovered after 5 h for the first G1 phase, after 12 h for the first S phase and after 29 h for the second S phase. To obtain cells in the second G1 phase, because of a partial desynchronization, the cells were recovered after 25 h and sorted by FACS with a FACSVantage according to their DNA content.

DNA combing. Combing was performed as described in ref. 3.

Received 23 April; accepted 1 July 2008.
Published online 17 August 2008.

1. Blow, J. J. & Dutta, A. Preventing re-replication of chromosomal DNA. *Nature Rev. Mol. Cell Biol.* **6,** 476–486 (2005).
2. Gilbert, D. M. Replication origin plasticity, Taylor-made: inhibition vs recruitment of origins under conditions of replication stress. *Chromosoma* **116,** 341–347 (2007).
3. Anglana, M., Apiou, F., Bensimon, A. & Debatisse, M. Dynamics of DNA replication in mammalian somatic cells: nucleotide pool modulates origin choice and interorigin spacing. *Cell* **114,** 385–394 (2003).
4. Debatisse, M., Berry, M. & Buttin, G. Stepwise isolation and properties of unstable Chinese hamster cell variants that overproduce adenylate deaminase. *Mol. Cell. Biol.* **2,** 1346–1353 (1982).
5. Toledo, F. *et al.* oriGNAI3: a narrow zone of preferential replication initiation in mammalian cells identified by 2D gel and competitive PCR replicon mapping techniques. *Nucleic Acids Res.* **26,** 2313–2321 (1998).
6. Toledo, F., Lachages, A. M., Mayau, V. & Debatisse, M. Initiation of DNA replication at the Chinese hamster origin *oriGNAI3* relies on local sequences and/or chromatin structures, but not on transcription of the nearby GNAI3 gene. *Nucleic Acids Res.* **27,** 1600–1608 (1999).
7. Svetlova, E. Y., Razin, S. V. & Debatisse, M. Mammalian recombination hot spot in a DNA loop anchorage region: A model for the study of common fragile sites. *J. Cell. Biochem.* **81,** 170–178 (2001).
8. DePamphilis, M. L. *et al.* Regulating the licensing of DNA replication origins in metazoa. *Curr. Opin. Cell Biol.* **18,** 231–239 (2006).
9. Berezney, R., Dubey, D. D. & Huberman, J. A. Heterogeneity of eukaryotic replicons, replicon clusters, and replication foci. *Chromosoma* **108,** 471–484 (2000).
10. Pasero, P., Bensimon, A. & Schwob, E. Single-molecule analysis reveals clustering and epigenetic regulation of replication origins at the yeast rDNA locus. *Genes Dev.* **16,** 2479–2484 (2002).
11. Lebofsky, R., Heilig, R., Sonnleitner, M., Weissenbach, J. & Bensimon, A. DNA replication origin interference increases the spacing between initiation events in human cells. *Mol. Biol. Cell* **17,** 5337–5345 (2006).
12. Buongiorno-Nardelli, M., Micheli, G., Carri, M. T. & Marilley, M. A relationship between replicon size and supercoiled loop domains in the eukaryotic genome. *Nature* **298,** 100–102 (1982).
13. Hancock, R. Internal organisation of the nucleus: assembly of compartments by macromolecular crowding and the nuclear matrix model. *Biol. Cell* **96,** 595–601 (2004).
14. Anachkova, B., Djeliova, V. & Russev, G. Nuclear matrix support of DNA replication. *J. Cell. Biochem.* **96,** 951–961 (2005).
15. Jenke, A. C. *et al.* Nuclear scaffold/matrix attached region modules linked to a transcription unit are sufficient for replication and maintenance of a mammalian episome. *Proc. Natl Acad. Sci. USA* **101,** 11322–11327 (2004).
16. Fernandez, M. A. *et al.* Matrix attachment regions and transcription units in a polygenic mammalian locus overlapping two isochores. *J. Cell. Biochem.* **67,** 541–551 (1997).
17. Razin, S. V. The nuclear matrix and chromosomal DNA loops: is their any correlation between partitioning of the genome into loops and functional domains? *Cell. Mol. Biol. Lett.* **6,** 59–69 (2001).
18. Pardoll, D. M., Vogelstein, B. & Coffey, D. S. A fixed site of DNA replication in eucaryotic cells. *Cell* **19,** 527–536 (1980).
19. Kitamura, E., Blow, J. J. & Tanaka, T. U. Live-cell imaging reveals replication of individual replicons in eukaryotic replication factories. *Cell* **125,** 1297–1308 (2006).
20. Lawlis, S. J., Keezer, S. M., Wu, J. R. & Gilbert, D. M. Chromosome architecture can dictate site-specific initiation of DNA replication in *Xenopus* egg extracts. *J. Cell Biol.* **135,** 1207–1218 (1996).
21. Lemaitre, J. M., Danis, E., Pasero, P., Vassetzky, Y. & Mechali, M. Mitotic remodeling of the replicon and chromosome structure. *Cell* **123,** 787–801 (2005).
22. Dietzel, S. & Belmont, A. S. Reproducible but dynamic positioning of DNA in chromosomes during mitosis. *Nature Cell Biol.* **3,** 767–770 (2001).
23. Wu, J. R. & Gilbert, D. M. A distinct G1 step required to specify the Chinese hamster DHFR replication origin. *Science* **271,** 1270–1272 (1996).
24. Li, F., Chen, J., Solessio, E. & Gilbert, D. M. Spatial distribution and specification of mammalian replication origins during G1 phase. *J. Cell Biol.* **161,** 257–266 (2003).
25. El Achkar, E., Gerbault-Seureau, M., Muleris, M., Dutrillaux, B. & Debatisse, M. Premature condensation induces breaks at the interface of early and late replicating chromosome bands bearing common fragile sites. *Proc. Natl Acad. Sci. USA* **102,** 18069–18074 (2005).
26. Gilbert, D. M., Miyazawa, H. & DePamphilis, M. L. Site-specific initiation of DNA replication in *Xenopus* egg extract requires nuclear structure. *Mol. Cell. Biol.* **15,** 2942–2954 (1995).
27. Merrick, C. J., Jackson, D. & Diffley, J. F. Visualization of altered replication dynamics after DNA damage in human cells. *J. Biol. Chem.* **279,** 20067–20075 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank E. Blackburn, R. Rothstein and F. Toledo for discussions and critical reading of the manuscript, and Genomic Vision for making available the DNA combing technology. S.C. is supported by a grant from the ARC (Association pour la Recherche sur le Cancer), and S.G. and N.A. are supported by a grant from the Ministère de la Recherche. The M.D. team is supported by La Ligue Nationale contre le Cancer and the Agence Nationale de la Recherche (ANR).

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.D. (michelle.debatisse@curie.fr).

naturejobs

THE CAREERS MAGAZINE FOR SCIENTISTS

We hear plenty about China's ascent as a science superpower, its new science institutes, ample government investment and large numbers of science students (see *Nature*'s recent China special at www.nature.com/news/specials/china/index.html). But China as a magnet for foreign talent is less well documented.

Chinese science graduates continue to go in droves to the United States and other countries for their postgraduate training. An analysis from the Commission on Professionals in Science and Technology in New York shows that by 2007 the two universities supplying the greatest number of graduates in all disciplines to US graduate-school programmes were Chinese — Tsinghua and Peking universities. The University of California, Berkeley, Seoul National University in South Korea, and the University of Michigan rounded out the top five.

But according to the Institute of International Education in New York, China is now also the fifth most popular destination for tertiary-level international students — behind the United States, the United Kingdom, France and Germany. The country's sheer size has something to do with it, with its expanding educational institutions, as does its prominence on the world stage. Students are realizing that some experience in China, not to mention some proficiency in Mandarin, is a nice thing to have on the CV. Chinese government programmes actively encourage foreign students to study there.

To import more graduate students, however, and not just undergraduates, China's institutions will have to address various challenges — especially in science training. As detailed in *Nature*'s China special and a recent *Naturejobs* feature (*Nature* **452,** 1028–1029; 2008), funding for graduate students, postdocs and researchers is poorly distributed and often inadequate. Simple things such as reagents are not as readily available as they could be. And there can be culture clashes.

But these things can be addressed, and when they are, China could become not only an exporter of graduate science talent but a major importer as well.

Gene Russo is editor of *Naturejobs*.

CONTACTS
Editor: Gene Russo

European Head Office, London
The Macmillan Building,
4 Crinan Street, London N1 9XW, UK
Tel: +44 (0) 20 7843 4961
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

European Sales Manager:
Andy Douglas (4975)
e-mail: a.douglas@nature.com
***Nature*events:**
Ghizlaine Ababou (+44 (0) 20 7014 4015)
e-mail: g.ababou@nature.com
UK Corporate:
Nils Moeller (4953)
Southwest UK/RoW:
Alexander Ranken (4944)
Northeast UK/Ireland:
Matthew Ward (+44 (0) 20 7014 4059)

France/Switzerland/Belgium:
Muriel Lestringuez (4994)
Scandinavia/Spain/Portugal/Italy:
Evelina Rubio-Hakansson (4973)
North Germany/The Netherlands/Eastern Europe: Reya Silao (4970)
South Germany/Austria:
Hildi Rowland (+44 (0) 20 7014 4084)

Advertising Production Manager:
Stephen Russell
To send materials use London address above.
Tel: +44 (0) 20 7843 4816
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com
***Naturejobs* web development:** Tom Hancock
***Naturejobs* online production:** Dennis Chu

US Head Office, New York
75 Varick Street, 9th Floor,
New York, NY 10013-1917
Tel: +1 800 989 7718
Fax: +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Manager: Peter Bless

India
Vikas Chawla (+91 1242881057)
e-mail: v.chawla@nature.com

Japan Head Office, Tokyo
Chiyoda Building, 2-37 Ichigayatamachi,
Shinjuku-ku, Tokyo 162-0843
Tel: +81 3 3267 8751
Fax: +81 3 3267 8746

Asia-Pacific Sales Manager:
Ayako Watanabe (+81 3 3267 8765)
e-mail: a.watanabe@natureasia.com
Business Development Manager, Greater China/Singapore:
Gloria To (+852 2811 7191)
e-mail: g.to@natureasia.com

NATURE|Vol 455|25 September 2008

MOVERS

Stephen Simpson, director of life sciences, Science Foundation Ireland, Dublin, Ireland

2000–08: Senior immunology editor, *Science*, Cambridge, UK
1998–2000: Senior research scientist, Edward Jenner Institute for Vaccine Research, Compton, UK
1997–98: Research fellow, St Vincent's University Hospital, Dublin, Ireland

As an undergraduate, Stephen Simpson read a *Scientific American* article on the immune system that captivated his imagination. He was intrigued, not only with the facts, but with the craft of communicating them to a broader audience. His curiosity would lead to a career in both science and editing. As the new director of life sciences at Science Foundation Ireland (SFI), Simpson is eager to help shape the research landscape of a country.

Simpson took his first degree in zoology at the University of Nottingham, UK. He moved to London to do his PhD at the National Institute for Medical Research, in part because of its pioneering use of transgenic expression in mice to probe immunology responses. Named a runner-up in *The Daily Telegraph*'s annual science-writing award, Simpson continued to develop his communication skills while breaking new ground in immunology.

Next came a postdoc with Cox Terhorst at Harvard Medical School, where Simpson explored the newly discovered immune system of the gut. As a result he developed productive mouse models of colitis and Crohn's disease. With Terhorst's encouragement, Simpson became a faculty instructor at Harvard for two more years.

In 1997, he crossed the Atlantic again, to St Vincent's University Hospital, Dublin. Despite Ireland's harsh funding climate at the time, he found an enthusiastic collaborator in Cliona O'Farrelly, who shared his interest in intestinal immunology. He next moved to the newly started Edward Jenner Institute for Vaccine Research in Newbury, UK, where he set up a small mucosal-immunology group to explore oral vaccines. But the shortage of independent positions concerned him. Then he spotted an advertisement for his dream job: immunology manuscript editor at *Science*.

"I didn't know if I would be competitive, but I had to put in an application," says Simpson. He got the job, and spent eight years there. Fellow *Science* editor Katrina Kelner says Simpson's extraordinary ability to engage others made him stand out. "His interest in science communication was apparent, as he helped develop podcasts as a new way to interact with the scientific community," says Kelner.

Simpson says his new role at SFI will allow him to use his editorial skills to shape the early-idea stage of the scientific process, by scrutinizing grant proposals rather than simply the final product. "We can't be too prescriptive, but we will surely encourage emerging areas of research, such as how gene regulation affects health and disease," he says. ■

Virginia Gewin

NETWORKS & SUPPORT

Two-body solution

Academic couples take heart. Your numbers, and your bargaining power, are rising, according to a report from the Michelle R. Clayman Institute for Gender Research at Stanford University. Dual academic couples now make up more than a third of the US professoriate. Universities should use them as a recruitment tool, the report concludes.

Dual-Career Academic Couples: What Universities Need to Know advises academic institutions that still practise secretive, inconsistent employment negotiations to join the growing number of universities with written policies or guidelines for partner hiring. Eighty-eight per cent of survey respondents who were hired as part of a couple said they would have gone elsewhere if their partner had not also been taken on, says Londa Schiebinger, director of the Clayman Institute. She led the survey, which included 9,000 full-time faculty members from 13 leading US research institutions.

Couple hiring could help institutes attract female faculty members, as 83% of women scientists in academic couples are partnered with another scientist. The University of Michigan in Ann Arbor is one of many that can demonstrate the recruitment benefits: in a recent study of its dual-career programme, 72% of respondents said the programme had made a critical difference in decisions to accept a position. "Our programme is tied to the heart of the institution — it is about recruiting and retaining faculty," says Glenda Haskell, assistant vice-provost of academic affairs.

Many candidates are concerned about when in the interview process they should bring up a partner's career, and often wait until they have a written offer, Schiebinger says. Beth Mitchneck, associate dean for academic affairs at the University of Arizona in Tucson, says: "Some interviewees fear that requesting a partner hire during the interview would preclude an offer." But she disagrees with that strategy. "Here we want to know as early as possible in order to best find satisfactory employment options," she says.

So how can a couple negotiate one offer into two positions? Schiebinger says the first step is for couples to discuss their individual career expectations by asking 'Whose career will we follow?'. They should then find out which universities are couple-friendly by looking for specific guidelines, typically found on the university provost's website. And, Schiebinger suggests, talk to couples who already work there. ■

Virginia Gewin

POSTDOC JOURNAL

Professional encounters

I guess it was time for some of my illusions to be shattered. I've had seven months of post-PhD life, in which I got the job I'd applied for, started fieldwork on a strange new species and continued the life of a happy researcher. While living in Ethiopia, I am exactly what I picture myself to be: a confident, intelligent, problem-solving field biologist. In America, however, a big surprise hit me — I am terrible at mingling with other academics.

I was completely unprepared for the vast difference between dealing with strangers in Ethiopia and strangers in academia. Frightening though the cultural and language gap may be, in Ethiopia any positive action or word on my part meets with approval. Mistakes are happily overlooked. A completely different fear grips me when I meet new people in academia. I'm expected to know the people I'm meeting, know what I'm talking about and know how my research intersects with that of the stranger in front of me. Or so it feels. I become completely flustered and feel far removed from the professional, successful scientist I want to be.

I am realizing that fieldwork is not preparing me for interacting with my scientific peers. That scares me a little. As I head back to the field, I wish I could have brought a self-help manual back to Ethiopia — something that could advise me on survival in the scientific rat race, without a single gelada to hide behind. ■

Aliza le Roux is a postdoctoral fellow in animal behaviour at the University of Michigan.

The Brown Revolution

No longer number two.

Norman Spinrad

Your Majesty; members of the Swedish Academy; fellow Nobel laureates; ladies, gentlemen and others, including former sceptics and tormentors — while convention and politesse would have me falsely protest my unworthiness to receive these two Nobel prizes for the same so-called discovery, receiving the Nobels for biology and for peace at the same time is hardly conducive to false modesty. Nor, after what I have had to endure in ridicule to finally stand here before you, am I in a particularly polite mood.

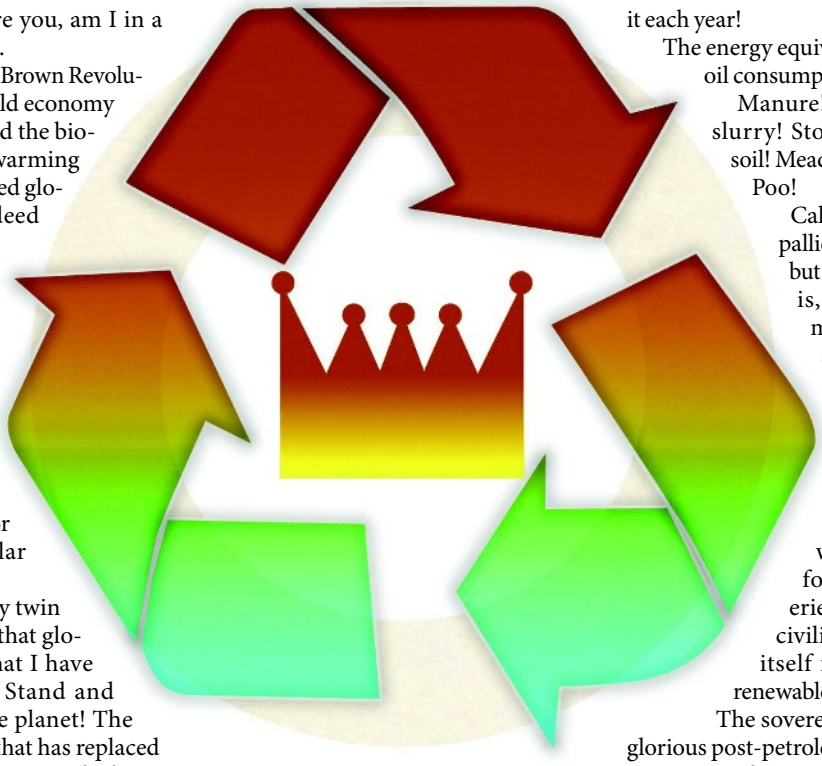
After all, even after the Brown Revolution has rescued the world economy from its energy crisis and the biosphere from the global warming crisis, that which has saved global civilization and indeed life on Earth itself, still cannot fly the public banner of its own true name. And I, who championed its cause, am still a victim of ridicule in odious cartoons and toilet graffiti, and will no doubt be even now denigrated as the winner of the Nobel Prize for Scatology in the popular yellow press.

Very well then! Let my twin Nobels combine under that glorious Brown Banner that I have borne to this victory! Stand and salute the saviour of the planet! The noble brown substance that has replaced petroleum and coal to become the basic abundant energy source of our triumphant global civilization!

Two hundred and fifty trillion kilos of it produced per annum by man and beast at the very least! Converted to over 45 trillion cubic metres of methane, the energy equivalent of more than the world's yearly oil production at its long bygone and unlamented peak! Free food as a by-product! Every year, forever, as long as the grasses grow and the rivers flow and it flows downhill through the food chain with them!

Captured from every retooled toilet in the world; from every modernized closed-circuit pigpen and cattle feed-lot, chicken coop and zoo! The methane brewed out of it with simple solar collectors and panels heating the great stills and generating abundant electricity to charge our hydrogen fuel cells, cars and trucks; to power our industry, light our homes and our cities!

And this methane, not burned in the open air, poisoning it with photochemical smog and raising the temperature of the planet with greenhouse gases, but confined to sealed, closed-loop generators, recycling the carbon dioxide from its combustion into carbohydrates to feed the hungry via sunlight, water and artificial photosynthesis; as the biomass of this planet has always done for just about as long as it has had one.



JACEY

Oh yes, after having championed its cause and saved the world from its own stupidity at the cost of making myself a laughing stock in the process, the only modesty that I am in any mood to proclaim here today, before this august company, is that I invented nothing that the biosphere wasn't doing quite well by itself until industrial civilization turned up its nose at the obviously natural and the naturally obvious, and broke the natural cycle.

Sunlight, water and carbon dioxide into carbohydrates via photosynthesis, carbohydrates fuelling animal life, animal life releasing carbon dioxide out one end and water and fecal fecundity out the others; and so it goes.

Yet we flushed it down our toilets; we dumped it into the seas; we let it pour into our rivers and pile up in feed-lot lagoons to spread disease and stink; we burned it in the open air and held our noses; we spent billions and billions throwing it away, or trying to; and all the while we were wasting the eternally renewable resource that has now, at last, become the energy source of global civilization, and allowed us to halt the greenhouse warming and feed the starving on People Chow.

Two hundred and fifty trillion kilos of it each year!

The energy equivalent of a peak year's oil consumption!

Manure! Dung! Crap! Fecal slurry! Stools! Excreta! Night soil! Meadow-muffins! Doody! Poo!

Call it by its hundreds of pallid euphemistic aliases, but call it also what it truly is, the long-despised, mocked, disdained substance we all produce each and every day of our lives, along with every animal we feed upon and many that we don't, that has become that which fuels the transformational machineries of the first human civilization to have freed itself from burning non-renewable fossil fuels.

The sovereign substance of our glorious post-petroleum age!

Here in this most hallowed hall of science and culture of our species let us finally speak it plain.

As you have crowned me with the golden wreath of the Nobel laureate, so do I now crown the Sun King of our glorious Brown Revolution with his true name before you!

Honour him!

Praise him!

Say it openly at long last!

All hail King Shit! ■

Norman Spinrad is the author of around 60 stories; more than 20 novels, including *Bug Jack Barron*, *Pictures at Eleven*, *Greenhouse Summer* and *The Druid King*; the classic *Star Trek* episode 'The Doomsday Machine'; and many other things. His most recent book is *Mexica*, the true story of Cortes's conquest of Mexico.

FUTURES